NEW SERIES, NO. 147 (VOL. XIX) SEPTEMBER, 1924

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

THE MEASUREMENT OF CORRELATION AND THE PROBLEM OF ESTIMATION[1]

By Frederick C. Mills, *Columbia University*

The two standard measures of correlation, the Pearsonian coefficient and the correlation ratio, have certain obvious limitations which materially restrict their usefulness. The former is limited to cases in which the relationship between two variables may be described by a straight line. The latter, while not restricted to linear functions, affords no basis for estimating the value of one variable from a given value of the other, since no equation of regression is derived in its computation. It is useful, moreover, only when based upon a large number of observations. There is need of a general measure, significant no matter what the functional relationship may be, and accompanied in all cases by a definite equation of relationship. It is the purpose of the present paper to present a method of computing such a general measure of correlation, a measure which is capable of computation directly from the values needed in fitting a curve by the method of least squares. It is suggested that this measure be termed the *index of correlation*, and be represented by the symbol ρ (Rho).[2]

The formula generally employed in computing the Pearsonian coefficient of correlation is

$$r = \frac{\Sigma(xy)}{N\sigma_x\sigma_y}.$$

It may be derived also from the equation

$$r^2 = 1 - \frac{S_y^2}{\sigma_y^2}$$

[1] This article presents a summary of the material contained in certain chapters of *Statistical Methods Applied to Economics and Business*, to be published by Henry Holt and Company.

[2] This symbol has been employed by Spearman to represent a coefficient of correlation based upon the squares of differences in rank. It has not been so widely used in this sense, however, that confusion should arise from its adoption for the general purpose here suggested.

It is of interest to note that Galton used the term "Index of Co-relation" for what is now called the coefficient of correlation.

where S_y^2 represents the mean square deviation about the line of regression of Y on X (*i. e.* $S=\sqrt{\frac{\Sigma(v^2)}{N}}$, v being the difference between each observed value of Y and the value as computed from the equation of regression). The quantity S_y is the standard error of estimate.

This last formula is generally applicable, with no limitation as to the type of curve which describes the relationship between the two variables. If the dispersion about the curve is less than the dispersion about the mean Y some degree of correlation may be assumed, and the relation between the two measures of dispersion constitutes a measure of this correlation. Employing the symbol suggested, we have, as a general formula for the index of correlation,[1]

$$\rho^2=1-\frac{S_y^2}{\sigma_y^2}.$$

This is a valid measure of correlation for any type of functional relationship. When the regression is linear, the values of ρ and r will be identical. When the curve employed to describe the relationship passes through the mean of each column, the values of ρ and η will be identical.

The facts set forth above are, of course, generally recognized, but such an index of correlation has not been widely employed, doubtless because no simple and standardized method of computing this index has been available. The calculation of S_y by measuring the deviation of each actual value of Y from the corresponding computed value is somewhat laborious, especially so since it must be preceded by the computations called for in fitting the desired curve. It is possible, however, to secure quite readily the value $\Sigma(v)^2$, and this fact makes practicable the employment of the general index of correlation proposed. A method of securing the value of S_y may be first explained.

Derivation of the Formula for the Standard Error of Estimate. Assume that, by the method of least squares, we have determined the constants in an equation of the type

$$Y_c=aW_1+bW_2+cW_3+dW_4.$$

(The symbol Y_c is used for a computed value of the dependent variable; the symbol Y represents an observed value. Throughout this exposition the capital letters represent deviations from the true origin, while small letters represent deviations from the arithmetic mean as origin.)

[1] With X dependent the general formula would be

$$\rho^2=1-\frac{S_x^2}{\sigma_x^2}.$$

For each residual, v, we have the relation

$$v = aW_1 + bW_2 + cW_3 + dW_4 - Y. \tag{1}$$

Multiplying throughout by v, and summing, we have

$$\Sigma(v^2) = a\Sigma(vW_1) + b\Sigma(vW_2) + c\Sigma(vW_3) + d\Sigma(vW_4) - \Sigma(vY). \tag{2}$$

But

$$\Sigma(vW_1) = 0^1$$
$$\Sigma(vW_2) = 0$$
$$\Sigma(vW_3) = 0$$
$$\Sigma(vW_4) = 0$$

therefore

$$\Sigma(v^2) = -\Sigma(vY). \tag{3}$$

Multiplying each equation (1) by Y, and adding, we have

$$\Sigma(vY) = a\Sigma(W_1Y) + b\Sigma(W_2Y) + c\Sigma(W_3Y) + d\Sigma(W_4Y) - \Sigma(Y^2). \tag{4}$$

Substituting in (3) the equivalent of $\Sigma(vY)$ we have

$$\Sigma(v^2) = \Sigma(Y)^2 - a\Sigma(W_1Y) - b\Sigma(W_2Y) - c\Sigma(W_3Y) - d\Sigma(W_4Y). \tag{5}$$

This gives us a means of obtaining the value $\Sigma(v^2)$ without computing the separate residuals.[2] In applying this to a particular case it is necessary only to replace W_1, W_2, W_3, W_4, etc., by the functions which actually appear as coefficients of the unknown quantities in the original equation. Thus in fitting a curve the equation to which is

$$Y = a + bX + cX^2 + dX^3$$

we find that

1 corresponds so W_1
X corresponds to W_2
X^2 corresponds to W_3
X^3 corresponds to W_4

[1] When the partial derivative of the expression

$$\Sigma(aW_1 + bW_2 + cW_3 + dW_4 - Y)^2 = \text{a minimum}$$

is equated to zero, with respect to the first unknown, a, we have

$$\Sigma[W_1(aW_1 + bW_2 + cW_3 + dW_4 - Y)] = 0.$$

Since

$$aW_1 + bW_2 + cW_3 + dW_4 - Y = v$$

we have as a necessary condition of fitting

$$\Sigma(vW_1) = 0.$$

When the partial derivative of the same expression with respect to b is equated to zero, we have

$$\Sigma[W_2(aW_1 + bW_2 + cW_3 + dW_4 - Y)] = 0$$

or, making the same substitution as in the preceding case,

$$\Sigma(vW_2) = 0.$$

Repeating the operation with respect to c and d we may show that

$$\Sigma(vW_3) = 0.$$

and

$$\Sigma(vW_4) = 0.$$

[2] This general method of determining the value of the sum of the squares of the residuals is explained in Brunt, D., *Combination of Observations*, Cambridge University Press, 1917, pp. 93–4.

Making these substitutions in equation (5) above, we have

$$\Sigma(v^2)=\Sigma(Y^2)-a\Sigma(Y)-b\Sigma(XY)-c\Sigma(X^2Y)-d\Sigma(X^3Y).$$

The quantity S_y is derived from the equation

$$S_y^2=\frac{\Sigma(v^2)}{N}\ ^1$$

where v is used to represent a deviation from a fitted curve. Accordingly, as a general expression for the standard error of estimate of Y, with W_1, W_2, W_3, W_4 . . . as independent variables, we have

$$S_y^2=\frac{\Sigma(Y^2)-a\Sigma(W_1Y)-b\Sigma(W_2Y)-c\Sigma(W_3Y)-d\Sigma(W_4Y)-\ .\ .\ .}{N} \quad (6)$$

As in the previous case, this may be applied to a particular problem by replacing W_1, W_2, W_3, W_4, etc., by the actual coefficients of the unknown quantities.

Derivation of the Formula for the Index of Correlation. We have adopted, as an index of the degree of correlation between two variables, the measure ρ, derived from the equation

$$\rho_{yx}^2=1-\frac{S_y^2}{\sigma_y^2}. \quad (7)$$

(The subscripts indicate that we are dealing with a single dependent variable, Y, and a single independent variable, X. With a single dependent variable, Y, and a number of independent variables, W_1, W_2, W_3, W_4, the expression might be written

$$\rho^2_{y\cdot w_1w_2w_3w_4}=1-\frac{S_y^2}{\sigma_y^2}.$$

Corresponding changes would be made in the subscripts for other changes in the symbols employed.)

Equation (7) above is equivalent to

$$\rho_{yx}^2=1-\frac{\Sigma(v^2)}{\Sigma(y^2)}, \quad (8)$$

where y represents a deviation from an origin at the mean of the y's.

[1] Since our object is to measure the actual "scatter" about the fitted curve, the formula $\frac{\Sigma(v^2)}{N}$ is used, rather than the formula $\frac{\Sigma(v^2)}{N-N_c}$ (where N represents the number of observations and N_c the number of constants in the equation to the fitted curve). The second formula would be used, in accordance with the theory of least squares, if we were seeking to determine the mean square error of an observation or of an observational equation.

But

$$\frac{\Sigma(y^2)}{N}=\frac{\Sigma(Y^2)}{N}-c_y^2$$

where Y represents the original values of the $Y-$variable and c_y represents the difference between the true origin and the mean of the Y's (the symbols c_y and c_x should not be confused with c, one of the constants in the equation to the fitted curve). Accordingly, we have

$$\rho_{yx}^2=1-\frac{\Sigma(v^2)}{\Sigma(Y^2)-Nc_y^2}. \tag{9}$$

But we have secured an expression for $\Sigma(v^2)$ which holds in the case of a curve fitted by the method of least squares. Taking first the general case, with a number of independent variables, and substituting the equivalent of $\Sigma(v^2)$ in the above equation, we have

$$\rho^2_{y\cdot w_1w_2w_3\ldots}=1-\frac{\Sigma(Y^2)-a\Sigma(W_1Y)-b\Sigma(W_2Y)-c\Sigma(W_3Y)-d\Sigma(W_4Y)-\ \ldots}{\Sigma(Y^2)-Nc^2y}. \tag{10}$$

Simplifying, we have, as a general formula for the index of correlation,

$$\rho^2_{y\cdot w_1w_2w_3\ldots}=\frac{a\Sigma(W_1Y)+b\Sigma(W_2Y)+c\Sigma(W_3Y)+d\Sigma(W_4Y)+\ \ldots\ -Nc^2y}{\Sigma(Y^2)-Nc^2y}. \tag{11}$$

Formula (11) holds in all cases in which the equation describing a functional relationship may be put in the form

$$Y_c=aW_1+bW_2+cW_3+dW_4+\ \ldots$$

This may be applied to a specific case by replacing W_1, W_2, W_3, W_4, etc., in the above formula by the functions which appear as coefficients of the unknown quantities in the original equation.

In the case of a straight line relationship between two variables, ρ is replaced by the symbol r, which represents the ordinary coefficient of correlation. As the general equation for r we have

$$r^2=\frac{a\Sigma(Y)+b\Sigma(XY)-Nc_y^2}{\Sigma(Y^2)-Nc_y^2}. \tag{12}$$

There are two special cases in which this formula may be simplified. If the origin be at the mean of the X's, we have

$$a=c_y=\frac{\Sigma(Y)}{N}$$

$$a^2=c_y^2=\frac{a\Sigma(Y)}{N}$$

$$Nc_y^2=a\Sigma(Y)$$

and the formula for r reduces to

$$r^2 = \frac{b\Sigma(xY)}{\Sigma(Y^2) - Nc_y^2}. \tag{13}$$

If the origin be at the mean of the Y's (it is not essential that it be also at the mean of the X's)

$$\Sigma(y) = 0, \text{ and } c_y = 0$$

and the formula for the coefficient of correlation becomes

$$r^2 = \frac{b\Sigma(Xy)}{\Sigma(y^2)}. \tag{14}$$

(In the latter case the general forumla for ρ may also be simplified by the elimination of the terms $a\Sigma(y)$ and Nc_y^2.)

The equivalence of equation (14) to the usual formula may be readily demonstrated. In the formula

$$r = \frac{\Sigma(xy)}{N\sigma_x\sigma_y}$$

x and y relate to deviations from an origin at the point of averages. When a straight line is fitted by the method of least squares in such a case, the constant b is determined from the relation

$$b = \frac{\Sigma(xy)}{\Sigma(x^2)}.$$

Under these conditions equation (14) may be written

$$r^2 = \frac{\dfrac{\Sigma(xy) \cdot \Sigma(xy)}{\Sigma(x^2)}}{\Sigma(y^2)}$$

which reduces directly to the usual product-moment formula.

The general formulae (6) and (11) above apply to cases of multiple as well as simple correlation. When the equation of regression is of the form

$$X_1 = a + b_{12 \cdot 34 \ldots n} X_2 + b_{13 \cdot 24 \ldots n} X_3 + b_{14 \cdot 23 \ldots n} X_4 + \ldots$$

we may write, for the standard error of estimate,

$$S_1^2 = \frac{\Sigma(X_1^2) - a\Sigma(X_1) - b_{12 \cdot 34 \ldots n}\Sigma(X_1X_2) - b_{13 \cdot 24 \ldots n}\Sigma(X_1X_3) - b_{14 \cdot 23 \ldots n}\Sigma(X_1X_4) - \ldots}{N} \tag{15}$$

which represents a simple adaptation of equation (6).

The coefficient of multiple correlation, R, may be computed from the formula

$$R^2 = 1 - \frac{S_1^2}{\sigma_1^2} \tag{16}$$

where S_1^2 represents the mean square deviation about the line of regression, and σ_1 represents the standard deviation of the dependent variable X_1. Equation (16) may be reduced to

$$R^2 = 1 - \frac{\Sigma(v^2)}{\Sigma(x_1^2)} \tag{17}$$

corresponding to equation (8) above. But $\Sigma(v^2)$, the sum of the squares of the deviations from the line of regression, may be computed from the equation

$$\Sigma(v^2) = \Sigma(X_1^2) - a\Sigma(X_1) - b_{12 \cdot 34} \ldots {}_n\Sigma(X_1X_2) - b_{13 \cdot 24} \ldots {}_n\Sigma(X_1X_3) - b_{14 \cdot 23} \ldots {}_n\Sigma(X_1X_4) - \ldots \tag{18}$$

which is derived in the same way as equation (5) above. Accordingly, equation (17) may be written

$$R^2 = 1 - \frac{\Sigma(X_1^2) - a\Sigma(X_1) - b_{12 \cdot 34} \ldots {}_n\Sigma(X_1X_2) - b_{13 \cdot 24} \ldots {}_n\Sigma(X_1X_3) - b_{14 \cdot 23} \ldots {}_n\Sigma(X_1X_4) - \ldots}{\Sigma(X_1^2) - Nc_1^2}. \tag{19}$$

(In this equation X_1 refers to deviations from the true origin, and c_1 represents the difference between the mean X_1 and the origin.)

Simplifying equation (19)

$$R^2 = \frac{a\Sigma(X_1) + b_{12 \cdot 34} \ldots {}_n\Sigma(X_1X_2) + b_{13 \cdot 24} \ldots {}_n\Sigma(X_1X_3) + b_{14 \cdot 23} \ldots {}_n\Sigma(X_1X_4) - Nc_1^2}{\Sigma(X_1^2) - Nc_1^2}. \tag{20}$$

This is a general formula for computing R. In the special case where the origin is at the point of averages, we have

$$R^2 = \frac{b_{12 \cdot 34} \ldots {}_n\Sigma(x_1x_2) + b_{13 \cdot 24} \ldots {}_n\Sigma(x_1x_3) + b_{14 \cdot 23} \ldots {}_n\Sigma(x_1x_4) + \ldots}{\Sigma(x_1^2)}. \tag{21}$$

Thus the coefficient of multiple correlation may be determined immediately from the values needed in fitting the line of regression by the method of least squares, with the single addition of $\Sigma(x_1^2)$.[1]

[1] A method of fitting such a line by the method of least squares is explained by H. R. Tolley and M. J. B. Ezekiel (*A Method of Handling Multiple Correlation Problems*, JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION, XVIII, 144, N. S., Dec., 1923). It is interesting to note that equation (21) is identical, after a slight change in form, with that derived by a different method by Tolley and Ezekiel.

The Calculation of the Index of Correlation. The computation of ρ for data being handled in individual pairs (*i. e.*, not classified in correlation tables) presents no new problem. The usual procedure in fitting curves by the method of least squares yields ρ as a by-product, with little extra calculation. The adequacy of the straight line fit may be tested quite readily by fitting a second degree parabola and comparing the results secured in the two cases.

The procedure is illustrated in the following example. It is desired to determine the relation between fluctuations in cotton production in the United States and in prices of middling upland cotton in New York. The original data and the first differences which are to be correlated are given in Table I.

The following values are necessary in fitting a straight line and a second degree parabola:

$N=$	22	$\Sigma(XY)=$	-117.3383
$\Sigma(X)=$	$-.36$	$\Sigma(X^2Y)=$	31.9956
$\Sigma(Y)=$	6.24	$\Sigma(X^2)=$	140.5634
$\Sigma(Y^2)=$	208.6688	$\Sigma(X^3)=$	-191.1285
$c_y=$	.2836	$\Sigma(X^4)=$	2288.1334

Solving for the constants in the linear equation we find that

$$a=.27$$
$$b=-.834$$

The equation to the best fitting straight line is, therefore,

$$Y=.27-.834X.$$

The coefficient of correlation is computed from the formula

$$r^2=\frac{a\Sigma(Y)+b\Sigma(XY)-Nc_y^2}{\Sigma(Y^2)-Nc_y^2}.$$

Its value is $-.687$.[1]

Their equation is

$$R^2=\frac{b_{12\cdot 34\ldots n}p_{12}+b_{13\cdot 24\ldots n}p_{13}+b_{14\cdot 23\ldots n}p_{14}+\ldots}{\sigma_1^2}$$

where p_{12}, etc., represent mean products of the type $\frac{\Sigma(x_1x_2)}{N}$

Employing these symbols, equation (18) above may be modified to furnish the following formula for the standard error of estimate:

$$S^2_{1\cdot 234\ldots n}=\sigma_1^2-b_{12\cdot 34\ldots n}p_{12}-b_{13\cdot 24\ldots n}p_{13}-b_{14\cdot 23\ldots n}p_{14}.$$

The results secured from this equation are identical with those secured from the usual formula for the standard deviation of order n:

$$\sigma^2_{1\cdot 234\ldots n}=\sigma_1^2(1-r^2_{12})(1-r^2_{13\cdot 2})(1-r^2_{14\cdot 23})\ldots(1-r^2_{1n\cdot 23\ldots n-1}).$$

[1] The sign of r cannot be determined from the general formula given. Its sign, however, will always be that of b, the constant measuring the slope.

TABLE I

COTTON PRODUCTION AND COTTON PRICES IN THE UNITED STATES, 1900–1923

(1) Crop year	(2) Production of cotton, excluding linters (in millions of bales)	(3) First differences, cotton production X	(4) Cotton prices, middling upland, N. Y. average, Sept. to May (cents per pound)	(5) Bradstreet's index (average, Sept. to May). Relative 1913–14 = 100	(6) Cotton prices, deflated (cents per pound)	(7) First differences, deflated cotton prices Y
1900–01	10.12	....	9.58	84.4	11.30	.....
1901–02	9.51	− .61	8.64	86.2	10.02	− 1.28
1902–03	10.63	+1.12	9.50	90.0	10.56	+ .54
1903–04	9.85	− .78	13.20	88.6	14.90	+ 4.34
1904–05	13.44	+3.59	8.69	89.3	9.73	− 5.17
1905–06	10.57	−2.87	11.40	92.3	12.35	+ 2.62
1906–07	13.27	+2.70	10.97	98.8	11.10	− 1.25
1907–08	11.11	−2.16	11.41	93.2	12.24	+ 1.14
1908–09	13.24	+2.13	9.81	91.3	10.74	− 1.50
1909–10	10.00	−3.24	14.62	100.6	14.53	+ 3.79
1910–11	11.61	+1.61	14.80	97.8	15.13	+ .60
1911–12	15.69	+4.08	10.34	100.0	10.34	− 4.79
1912–13	13.70	−1.99	12.35	104.8	11.78	+ 1.44
1913–14	14.16	+ .46	13.40	100.0	13.40	+ 1.62
1914–15	16.13	+1.97	8.63	105.2	8.20	− 5.20
1915–16	11.19	−4.94	12.04	121.2	9.93	+ 1.73
1916–17	11.45	+ .26	18.29	151.0	12.11	+ 2.18
1917–18	11.30	− .15	29.96	197.9	15.14	+ 3.03
1918–19	12.04	+ .74	30.06	203.1	14.80	− .34
1919–20	11.42	− .62	38.63	226.3	17.07	+ 2.27
1920–21	13.44	+2.02	16.90	152.9	11.05	− 6.02
1921–22	7.95	−5.49	18.67	127.2	14.68	+ 3.63
1922–23	9.76	+1.81	26.26	149.7	17.54	+ 2.86
		−22.85				−25.55
		+22.49				+31.79
		− .36				+ 6.24

Solving now for the constants in the equation to a second degree parabola, we have

$$a = \ 1.1966$$
$$b = -1.03$$
$$c = - \ .1455$$

The desired equation is

$$Y = 1.1966 - 1.03X - .1455X^2.$$

The index of correlation is computed from the formula

$$\rho^2 = \frac{a\Sigma(Y) + b\Sigma(XY) + c\Sigma(X^2Y) - Nc_y^2}{\Sigma(Y^2) - Nc_y^2}.$$

The value of ρ is found to be .742 (no positive or negative sign attaches to ρ unless the regression be linear). These values indicate that estimates based upon the higher equation would, in general, be closer to the truth than those based upon the linear equation.[1] By a simple

[1] One should have the probable error of the difference between r and ρ (or between r^2 and ρ^2) to know what constitutes a significant difference between these values. The formula for this probable error has not been derived.

modification of these methods curves of higher degree or of other types may be fitted and the index ρ computed.

The procedure is essentially the same in measuring the relationship between two variables when the data are arranged in the form of a correlation table. Such a table does not lend itself readily to the computation of the higher product-sums (as ΣX^2Y, etc.) which are needed in fitting curves of higher degree. For this purpose a rearrangement of the data in columnar form is advisable. In fitting such curves an arbitrary origin may be selected and the class-interval unit employed throughout, just as is done in following the usual method of measuring linear correlation.

In handling grouped data, it may be noted, the system here proposed merely provides a method of estimating the mean of an array. Unless the variability of all the arrays be equal (*i. e.* unless the system be homoscedastic) it cannot be assumed that the general measure of scatter about the line of regression will serve as a measure of the variability of each array. It is not necessary in using an equation of regression that a homoscedastic distribution be assumed, though the usefulness of the measures secured would be much greater in cases where such a distribution prevailed.[1]

Tests for Lines of Regression. The adequacy of a straight line to describe the relationship between two variables, when the data are arranged in the form of a correlation table, may be tested by the measure

$$\zeta = \eta^2 - r^2.$$

When the straight line of best fit passes through the mean of each array the values of η and r will be identical, and the equation to the line of regression may be used with confidence. By the use of Blakeman's formula for the probable error of ζ it is possible to determine whether a given difference between η^2 and r^2 is due to a mere chance fluctuation of sampling, or represents a fundamental departure from linearity.[2]

It is obvious that a measure similar to ζ and applicable whenever we are dealing with grouped data may be derived with reference to the index of correlation, ρ. That is, we may employ the value

$$\eta^2 - \rho^2$$

as a test of the adequacy of any specific curve, fitted by the method of least squares, to describe the relationship between two variables, pro-

[1] *Cf.* Karl Pearson: Mathematical Contributions to the Theory of Evolution, XIV. "On the General Theory of Skew Correlation and Non-Linear Regression." *Drapers' Company Research Memoirs, Biometric Series II*, 1905, pp. 21 *et seq.*

[2] "On Tests for Linearity of Regression in Frequency Distributions." John Blakeman. *Biometrika*, Vol. 4, pp. 332–350.

vided always that the data permit η to be computed. To interpret this value, of course, it is always necessary to know the type of curve to which ρ refers. If a certain curve passes through the mean of each array, ρ, as computed for that curve, and η will be identical in value. To determine whether a given difference is or is not significant, it would be necessary to know the probable error of the difference between η^2 and ρ^2. This formula has not been derived.

Tests analogous to this were worked out by Karl Pearson for second and third degree parabolas, the computations involved in fitting being based upon the method of moments in his examples.[1] It has not been established that Pearson's conditions coincide with the somewhat simpler condition

$$\eta^2-\rho^2=0$$

which is here suggested, but in comparing results secured from the different methods in a specific problem no significant differences were revealed.

The Measurement of Errors of Estimate. The value of the technique of correlation derives in large part from its practical utility in connection with the problem of estimation. To the economic and business statistician equations of regression, measures of standard error of estimate and coefficients of correlation are of interest primarily because of their bearing upon the problems of determining probable production, probable price, probable business changes. Nor is the making of estimates confined to attempts to forecast future changes. Whenever we seek to determine the most probable value from a number of different observations, or whenever we employ an equation which describes the average relationship between two or more variables we are engaged in estimating. It is essential that every such estimate be accompanied by an appropriate measure of the probable error involved. The measurement of correlation cannot be discussed apart from this central problem concerned with the measurement of errors of estimate.

In considering such measures in the preceding pages we have been concerned with *arithmetic* values throughout. In dealing with an equation of the type

$$Y_c=a+bX+cX^2+ \quad . \quad . \quad .$$

each residual may be represented by the expression

$$v=a+bX+cX^2+ \quad . \quad . \quad . \quad -Y.$$

[1] *Drapers' Company Research Memoirs, Biometric Series II.* Mathematical Contributions to the Theory of Evolution, XIV. "On the General Theory of Skew Correlation and Non-Linear Regression."

The residuals, as thus computed, are in terms of natural numbers. The standard error of estimate, as derived from the equation

$$S_y^2 = \frac{\Sigma(Y^2) - a\Sigma(Y) - b\Sigma(XY) - c\Sigma(X^2Y) - \quad . \quad . \quad .}{N},$$

is also expressed in these terms. The standard error of estimate, computed after the method outlined above, is always in the same terms as the dependent variable. If this variable be in terms of logarithms or reciprocals the measure of the errors of estimate is correspondingly changed in form, and this change must be taken account of in the interpretation of S and of the index of correlation based upon this value. Certain interesting questions are raised in this interpretation. These questions may best be considered in terms of a specific problem.

In Table II the production and price of oats in the United States from 1881 to 1913 are recorded. Appropriate lines of trend were

TABLE II

PRODUCTION AND PRICE OF OATS IN THE UNITED STATES, 1881–1913*

Year	Production of oats in U. S. (millions of bu.)	Straight line trend of production†	Ratio of actual production to trend value X	Price of oats in Chicago (cents per bu.)	Straight line trend of price‡	Ratio of actual price to trend value Y
1881	416	448	.929	47	36.0	1.30
1882	488	471	1.036	37	35.3	1.05
1883	571	494	1.156	31	34.6	.90
1884	583	517	1.128	29	34.0	.85
1885	629	540	1.165	28	33.2	.84
1886	624	563	1.108	25	32.5	.77
1887	659	586	1.124	30	31.2	.96
1888	701	609	1.151	24	30.5	.79
1889	751	632	1.188	24	29.8	.81
1890	523	655	.798	43	29.0	1.48
1891	738	678	1.088	31	28.3	1.10
1892	661	701	.943	30	27.5	1.09
1893	639	724	.882	31	26.8	1.16
1894	662	747	.886	28	26.1	1.07
1895	824	770	1.070	19	25.3	.75
1896	780	793	.983	18	23.6	.76
1897	791	816	.969	24	25.0	.96
1898	843	839	1.005	25	26.4	.95
1899	926	862	1.074	23	27.8	.83
1900	914	885	1.033	25	29.2	.86
1901	778	908	.857	42	30.6	1.37
1902	1053	931	1.131	33	32.0	1.03
1903	869	954	.911	38	33.4	1.14
1904	1009	977	1.033	30	34.8	.86
1905	1090	1000	1.090	31	36.2	.86
1906	1036	1023	1.013	39	37.6	1.04
1907	805	1046	.770	51	38.0	1.31
1908	851	1069	.796	52	40.4	1.29
1909	1068	1092	.978	43	41.8	1.03
1910	1186	1115	1.064	35	43.2	.81
1911	922	1138	.810	51	44.6	1.14
1912	1418	1161	1.221	37	46.0	.80
1913	1122	1184	.948	41	47.4	.87

*For the use of these data I am indebted to Mr. H. B. Killough, of the Bureau of Agricultural Economics, U. S. Department of Agriculture.

†This line of trend was fitted to data covering a longer period than that covered in the present study.

‡The entire period has been broken into two parts, 1881 to 1895 and 1896 to 1913. A straight line of trend has been fitted to the data of each period.

fitted to these series and the ratios of the actual values of each series to the trend values determined. It is desired to determine the relation between these two variables. A hyperbolic curve of the general type $Y = aX^b$ appears to be an appropriate form to employ in describing such a relationship. To fit this curve by the method of least squares, the equation must be reduced to the logarithmic form

$$\text{Log } Y = \log a + b \log X.$$

Computing the values required in fitting a curve of this type, we have

$$N = 33$$

$$\Sigma(\log Y) = -.328500 \qquad \Sigma(\log X \cdot \log Y) = .-1143005$$

$$\Sigma(\log X) = .037535 \qquad \Sigma(\log^2 X) = .096423$$

Solving the necessary normal equations, we find that

$$\log a = -.00861$$

$$b = -1.182059$$

The required equation is

$$\text{Log } Y = (9.99139 - 10) - 1.182059 \log X$$

or

$$Y = .9804X^{-1.182059}.$$

This is the equation which describes the average relationship between the production and the price of oats (when the actual figures for each are expressed as ratios to the respective lines of trend). The corresponding curve is plotted in Figure II.

A Logarithmic Measure of Errors of Estimate. The next step is the determination of the reliability of this equation as a means of estimating the price of oats from a knowledge of production. For this purpose we must compute the standard error of estimate, S. Since the fitting process was carried through in terms of logarithms, the standard error may be computed in the same terms. Following the procedure explained in earlier sections, we may derive the following equation relating to the logarithmic curve just fitted:

$$S^2_{\log y} = \frac{\Sigma(\log^2 Y) - \log a \Sigma(\log Y) - b\Sigma(\log X \cdot \log Y)}{N}.$$

Substituting the proper values, we have

$$S^2_{\log y} = \frac{.21721807 - (-.00861 \times -.32850) - (-1.182059 \times -.1143005)}{33}$$

$$= \frac{.07927976}{33}$$

$$= .002402417$$

$$S_{\log y} = .04901$$

The standard error of estimate, in terms of logarithms, is .04901. As long as we deal with logarithms, this is to be interpreted precisely as is the standard error with respect to other curves. Assuming a normal distribution of logarithms about the curve which describes the average relationship, the chances are 68 out of 100 that the logarithm of a given estimate will not differ from the logarithm of the actual value by more than .04901. This means, simply, that we are dealing throughout in terms of ratios instead of absolute figures. The difference between the logarithms of two numbers is the logarithm of the ratio of one of the original numbers to the other. Thus the absolute value of S in a given case will depend upon the magnitude of the values with which we are dealing. If the user desires to reduce S to absolute values, it must be done always with reference to a given estimate.

The "zones of estimate" marked out with respect to a logarithmic curve will differ materially from similar zones relating to simple linear equations. In the simple case a zone extending $1S$ on each side of the estimating equation has the same absolute width throughout its length, and centers always at the line of regression. The logarithmic zone, when measured in natural numbers, is of varying width, and, moreover, is not of the same width on each side of the plotted curve. It is true, however, that the ratios on the two sides of the curve are always equal. That is, the ratio of a value $1S$ less than the computed value to the computed value is the same as the ratio of the latter to a value $1S$ greater. And when the curves are plotted on paper ruled logarithmically the zone included within a distance $1S$ on each side of the plotted curve is reduced to the symmetry and evenness found in the simpler case. A person accustomed to thinking in terms of ratios and to the use of logarithmic paper can readily interpret this measure.

Since the ratios are equal throughout, the standard error of estimate may be expressed in terms of ratios. In the present example we have

$$S_r = \text{anti-log } S_{\log y} = \text{anti-log } .04901 = 1.12$$

where the symbol S_r is used to represent the standard error of estimate in terms of ratios. $S_{\log y}$, as derived above, is positive, hence the ratio exceeds unity. It is the ratio of the larger number to the smaller. It represents the upper limit of the zone of estimate. The lower limit, in terms of ratios, may be secured by putting $S_{\log y}$ in negative terms, and getting the corresponding natural value. The value $-.04901 = 9.95091 - 10$, which is the logarithm of .8933. In this form the ratio is based upon the relation of the smaller to the larger number. To make S_r readily intelligible we may combine the two, writing

$$S_r = .89 \text{ to } 1.12.$$

Interpreting this, it means that in 68 cases out of 100 the actual value will not be less than 89 per cent of the estimate, or more than 112 per cent of the estimate. This has a simple, definite meaning, more significant for most practical purposes than a similar measure in terms of absolute values.[1]

To find the values of $2S$ and $3S$, the value of $S_{\log y}$ must be multiplied by 2 and 3, and the resulting values reduced to natural numbers. Summarizing the results, we have

$$S_r = .89 \text{ to } 1.12$$
$$2S_r = .80 \text{ to } 1.25$$
$$3S_r = .71 \text{ to } 1.40$$

The values given for S_r indicate the probable percentage limits within which actual value and estimated value should fall in 68 out of 100 cases. The values given for $2S$ indicate the probable percentage limits in 95 out of 100 cases. The values of $3S_r$ indicate the probable percentage limits in 99.7 cases out of 100, always on the assumption of a normal distribution of the logarithms of the actual values about the fitted line.

The Index of Correlation based on Logarithmic Measures. We have still to compute the third measure—the abstract index of correlation. For an equation of the type

$$\log Y = \log a + b \log X$$

the formula for ρ reduces to

$$\rho^2 = \frac{\log a \Sigma(\log Y) + b\Sigma(\log X \cdot \log Y) - Nc^2_{\log y}}{\Sigma(\log^2 Y) - Nc^2_{\log y}}$$

where $c_{\log y}$ represents the difference between the arithmetic mean of the logarithms of the Y values and the origin (in this case, zero on the logarithmic scale). Substituting the proper values we have

$$\rho^2 = \frac{(-.00861 \times -.32850) + (-1.182059 \times -.1143005) - (33 \times .00009909}{.21721807 - (33 \times .00009909)}$$

$$= \frac{.13466835}{.2139481}$$

$$= .629444$$

$$\rho = .793$$

[1] The significance of a measure of errors of estimate in percentage form was pointed out by D. H. Davenport in 1922, in an unpublished article, and such a measure has been employed in several studies. There has not been available, however, a ready method of computing this measure, and its possibilities have not, therefore, been fully realized.

The significance of this measure may be clearer if viewed in terms of the relationship

$$\rho^2 = 1 - \frac{S^2_{\log y}}{\sigma^2_{\log y}}.$$

In the present case these values are

$$S_{\log y} = .04901$$
$$\sigma_{\log y} = .08052$$

Inserting the squared values in the formula, we have

$$\rho^2 = 1 - \frac{.002402417}{.00648328}$$
$$= .629444$$
$$\rho = .793^1$$

What does this value measure? The value of ρ, as we have seen, depends upon the relation between two quantities, S^2 and σ^2, the former measuring the variability about the fitted line, the latter measuring the variability about the mean of the dependent variable. In the example given in the early pages of this paper, the variability was measured in absolute terms, and the value of ρ depended upon the relation between the two given measures of *absolute variability*. The sole difference in the present case is that we are working in terms of *logarithmic variability*, or *ratio variability*, deviations being measured in terms of logarithms, or of ratios. The value of ρ now depends upon the relation between the *ratio variability* about the fitted curve and the *ratio variability* about the geometric mean of the Y's. (It is the geometric mean of the Y's, because that is the value corresponding to the arithmetic mean of the Y logarithms. The value $\sigma^2_{\log y}$ is the mean square logarithmic deviation about the arithmetic mean of the logarithms of the Y-values.)

We have here a set of measures, therefore, which perform in the field of ratios precisely the same service as is performed in the field of natural numbers by S and ρ (in the linear case, r). These measures are secured in precisely the same way as are S and ρ, except that the original equation of relationship is one in which the dependent variable is log Y (or, in the reverse case, log X). The general formulae for getting these values are the same as in dealing with natural numbers, except that log Y replaces Y throughout. It is precisely the same operation as that of using logarithmic or ratio paper instead of natural scale paper.

[1] The symbol ρ is used for this measure of correlation, instead of r, even though the relationship in logarithmic form is linear. This is done because such a measure, in terms of logarithms, cannot be interpreted in precisely the same way as the ordinary coefficient of correlation.

The Measurement of Relationship in Terms of Reciprocals. Another type of curve may be used to describe the relationship between the production and price of oats, and its use introduces us to a third field of correlation, a field in which somewhat new concepts enter and in which the various measures must be interpreted in still another way. This is a curve of the type

$$Y = \frac{1}{a+bX}$$

which may be expanded by adding additional terms to the denominator, as

$$Y = \frac{1}{a+bX+cX^2}.$$

This hyperbolic form has been used in several studies as an approximation to a "demand" curve for various commodities.

The equation to a curve of this type may be written

$$\frac{1}{Y} = a+bX$$

which is the equation to a straight line describing the relationship between the reciprocals of the Y's and the original $X-$ values.

The values needed in solving for the constants in this equation, by the method of least squares, and in computing the index of correlation, are

$$N = 33 \qquad \Sigma\left(\frac{X}{Y}\right) = 35.2571485$$

$$\Sigma\left(\frac{1}{Y}\right) = 34.3360320 \qquad \Sigma(X^2) = 34.168554$$

$$\Sigma(X) = 33.338 \qquad \Sigma\left(\frac{1}{Y}\right)^2 = 36.8570294$$

Solving, we have

$$a = -.1357$$
$$b = \ 1.1643$$

The desired equation is, therefore,

$$\frac{1}{Y} = -.1357+1.1643X.$$

To determine the utility of this equation we must have the standard error and the index of correlation. The two necessary formulae may be derived as in the preceding cases. For $S_{\frac{1}{y}}^2$, we have

$$S_{\frac{1}{y}}^2 = \frac{\Sigma\left(\frac{1}{Y}\right)^2 - a\Sigma\left(\frac{1}{Y}\right) - b\Sigma\left(\frac{X}{Y}\right)}{N}.$$

Inserting this value of $S_{\frac{1}{y}}^2$ in the general formula for the index of correlation

$$\rho^2 = 1 - \frac{S_{\frac{1}{y}}^2}{\sigma_{\frac{1}{y}}^2}$$

and simplifying, we have

$$\rho^2 = \frac{a\Sigma\left(\frac{1}{Y}\right) + b\Sigma\left(\frac{X}{Y}\right) - Nc_{\frac{1}{y}}^2}{\Sigma\left(\frac{1}{Y}\right)^2 - Nc_{\frac{1}{y}}^2}.$$

Inserting the proper values in these two equations, and solving, we secure

$$S_{\frac{1}{y}} = .1191$$

$$\rho = .766$$

For the standard deviation of the original Y-values, in terms of reciprocals, we secure

$$\sigma_{\frac{1}{y}} = .1851.$$

(The subscript $\frac{1}{y}$ is used in connection with each of these measures, as they should be distinguished from measures based upon natural numbers or logarithms.)

How may we interpret these results? As in all other problems of this type the equation gives us a means of estimating Y from a known value of X. The standard error $S_{\frac{1}{y}}$ serves as a measure of the reliability of such estimates, and ρ is an abstract measure of the degree of relationship between the two variables. But in the present case all these measures are in terms of reciprocals. The equation enables us to estimate the reciprocal of Y, the standard error has significance only in the form of a reciprocal, and ρ depends upon the relation between two measures ($S_{\frac{1}{y}}^2$ and $\sigma^2_{\frac{1}{y}}$) both of which are in terms of reciprocals.

In making use of the standard error of estimate in this form, the limits of the "zones of estimate" must first be computed by securing the estimated values of the Y-variable in reciprocal form, plus and minus the quantities $S_{\frac{1}{y}}$, $2S_{\frac{1}{y}}$, and $3S_{\frac{1}{y}}$. The reciprocals thus secured may be reduced to natural numbers to give the limits of estimate in more useful form.

Just as with logarithms the value in absolute terms of a given difference between reciprocals varies at different points within the range of

the Y-values. Accordingly, the limits of reliability determined from $S_{\frac{1}{y}}$ should be expressed in natural numbers only after a particular estimate has been made.

In interpreting ρ similar considerations enter. When the reciprocals of the dependent variable are employed in fitting a curve, variability about the fitted line is measured in terms of reciprocals, and the variability of the original series is measured in the same terms. That is, $\sigma_{\frac{1}{y}}$ is computed from the differences between the reciprocals of the actual values and the arithmetic mean of all these reciprocals. But the arithmetic mean of these reciprocals is the reciprocal of the harmonic mean. Thus, in effect, the value of this index of correlation depends upon the relation between variation about the fitted line and variation about the *harmonic mean* of the dependent series, variation in both cases being measured in terms of reciprocals.

Arithmetic, Geometric and Harmonic Measures. We have, therefore, three broad families of curves for describing the relationship between variable quantities. These are

1. Curves in the fitting of which natural values of the dependent variable are employed. Equations to all curves of this family will be of the type

$$Y=f(X).$$

2. Curves in the fitting of which logarithms of the dependent variable are employed. In all such cases the equations will be of the type

$$\text{Log } Y=f(X).$$

3. Curves in the fitting of which reciprocals of the dependent variable are employed. For these curves the equations will be of the type

$$\frac{1}{Y}=f(X).$$

In any one of these three cases the equations may be linear or non-linear. In so far as this problem of interpretation is concerned, there is no limitation as to the function of X which may be employed. (The computation of S and ρ by the methods suggested above is limited, as has been pointed out, to cases to which the method of least squares is applicable.)

The standard error of estimate for the first family of curves is derived in terms of the original units of measurement (for the dependent variable) and has a direct and simple meaning in these terms. The index of correlation, for curves of this type, is a measure of the degree

to which the *absolute variability* of the dependent variable may be lessened by measuring deviations from the fitted curve instead of from the *arithmetic mean.*

The standard error of estimate for the second family of curves is derived, by the method outlined, in terms of logarithms. It is more convenient in general to give it meaning in terms of *ratios.* The index of correlation, $\rho_{\log y}$ is a measure of the degree to which the *logarithmic or ratio variability* of the dependent variable may be lessened by computing deviations (or ratios) with respect to the fitted curve instead of the *geometric mean.*

The standard error of estimate for the third family of curves is derived by the same process as in the other cases, but emerges as a *reciprocal.* The index of correlation, $\rho_{\frac{1}{y}}$, is a measure of the degree to which the variability of the dependent variable, *in terms of reciprocals,* may be lessened by computing reciprocal deviations from the fitted curve instead of from the *harmonic mean.*

It is clear, therefore, that the choice of a type of curve to describe a given relationship must be governed by basic considerations as to the type of average which is most appropriate as a measure of the central tendency of the given distribution. And this brings in a related question as to whether the dispersion about this average more nearly approximates the normal type when measured in absolute terms, in logarithms, or in reciprocals. In selecting a curve and in using the measures S and ρ there is always present an implicit assumption with respect to these points.

A Comparison of Arithmetic, Geometric and Harmonic Measures. The contrast between these different methods may be brought home most effectively by comparing the results obtained when curves of these three types are fitted to the same data. The computations involved in fitting curves of the second and third types (logarithmic and

RELATION BETWEEN THE PRODUCTION AND PRICE OF OATS, 1881–1913
COMPARISON OF RESULTS OF CURVE FITTING
(Prices are the dependent variable in each case)

	Equation	Standard error of estimate	Index of correlation
A	$Y = 2.24 - 1.236\,X$	$S_y = .12$	$r = .783$
B	$\text{Log } Y = -.00861 - 1.182059 \log X$	$S_{\log y} = .04901$	$\rho_{\log y} = .793$
C	$\frac{1}{Y} = -.1357 + 1.1643\,X$	$S_{\frac{1}{y}} = .1191$	$\rho_{\frac{1}{y}} = .766$

reciprocal) have been illustrated with reference to the data of oat production and prices (Table II). A straight line (arithmetic) is fitted to the same data, and the necessary accompanying measures computed. The three sets of results are brought together in the summary as shown by the table on page 292.

It is impossible to compare the three standard errors as they stand, since only the first one is in the original units of measurement (ratio of actual price to normal). In the following table are given estimates, based on each of these equations, as to the most probable price (in terms of ratio to normal) which would accompany each of five different conditions of production. Each estimate is accompanied by a series of values which indicate the limits set by the standard error. Throughout the values of the estimates plus and minus S, $2S$ and $3S$, are given in order to indicate the probable scatter of actual values about the estimates. The different amounts of variation which may be expected about each of the three lines of relationship are measured by the actual differences between the estimates and the limiting cases. These differences are given in the columns headed Δ. All values in this table are comparable, being reduced to the original units (ratio of price to normal).

A careful study of the table on page 294 should make clear the nature of estimates based on the three types of equations here presented. The fundamental differences lie not so much in the actual values of the estimates, as in the standard errors which measure the reliability of these estimates and indicate the limits within which the actual values are likely to fall. In other words, the differences lie in the assumptions made as to the character of the "scatter" about the curves.

The measure S_y, which relates to the arithmetic curve, gives the same absolute range to errors of estimate whether the estimated value be high or low. An arithmetic dispersion about the curve is assumed. In each case the estimate is the arithmetic mean of the value which exceeds the estimate by an amount equal to S_y (or any multiple of S_y) and the value which falls below it by an equal amount. These conditions are brought out graphically in Figure I. The original points are plotted, the straight line of relationship (arithmetic) is shown, and zones of estimate having widths, respectively, of $2S$, $4S$ and $6S$, centering at the fitted line, are marked out.

The measure $S_{\log y}$ gives the same relative or percentage range to errors of estimate, whether the estimate be high or low. This means that the absolute range within which the actual values should fall is much less when the estimates are low than when they are high. It assumes a geometric dispersion about the curve which describes the

TABLE III

COMPARISON OF PRICE ESTIMATES AND OF STANDARD ERRORS OF ESTIMATE BASED ON THREE EQUATIONS RELATING TO THE PRODUCTION AND PRICE OF OATS

(1) Value of X (ratio of production to normal)	(2) Estimated value of Y (ratio of price to normal) from arithmetic equation (A)	(3) Limits of arithmetic estimate	(4) Δ	(5) Estimated value of Y from logarithmic equation (B)	(6) Limits of logarithmic estimate	(7) Δ	(8) Estimated value of Y from reciprocal equation (C)	(9) Limits of estimate; reciprocal	(10) Δ
.5	1.622	Estimate $+3S$ = 1.982	+.36	2.224	Estimate $+3S$ = 3.114	+.890	2.240	Estimate $+3S$ = 11.223	+8.983
		" $+2S$ = 1.862	+.24		" $+2S$ = 2.780	+.556		" $+2S$ = 4.803	+2.563
		" $+S$ = 1.742	+.12		" $+S$ = 2.491	+.267		" $+S$ = 3.055	+ .815
		" $-S$ = 1.502	−.12		" $-S$ = 1.979	−.245		" $-S$ = 1.768	− .472
		" $-2S$ = 1.382	−.24		" $-2S$ = 1.779	−.445		" $-2S$ = 1.461	− .779
		" $-3S$ = 1.262	−.36		" $-3S$ = 1.579	−.645		" $-3S$ = 1.244	− .996
.8	1.251	Estimate $+3S$ = 1.611	+.36	1.276	Estimate $+3S$ = 1.786	+.510	1.257	Estimate $+3S$ = 2.281	+1.024
		" $+2S$ = 1.491	+.24		" $+2S$ = 1.595	+.319		" $+2S$ = 1.794	+ .537
		" $+S$ = 1.371	+.12		" $+S$ = 1.429	+.153		" $+S$ = 1.478	+ .221
		" $-S$ = 1.131	−.12		" $-S$ = 1.136	−.140		" $-S$ = 1.093	− .164
		" $-2S$ = 1.011	−.24		" $-2S$ = 1.021	−.255		" $-2S$ = .967	− .290
		" $-3S$ = .891	−.36		" $-3S$ = .906	−.370		" $-3S$ = .867	− .390
1.0	1.004	Estimate $+3S$ = 1.364	+.36	.980	Estimate $+3S$ = 1.372	+.392	.972	Estimate $+3S$ = 1.490	+ .518
		" $+2S$ = 1.244	+.24		" $+2S$ = 1.225	+.245		" $+2S$ = 1.265	+ .293
		" $+S$ = 1.124	+.12		" $+S$ = 1.098	+.118		" $+S$ = 1.100	+ .128
		" $-S$ = .884	−.12		" $-S$ = .872	−.108		" $-S$ = .871	− .101
		" $-2S$ = .764	−.24		" $-2S$ = .784	−.196		" $-2S$ = .789	− .183
		" $-3S$ = .644	−.36		" $-3S$ = .696	−.284		" $-3S$ = .722	− .250
1.2	.757	Estimate $+3S$ = 1.117	+.36	.790	Estimate $+3S$ = 1.106	+.316	.793	Estimate $+3S$ = 1.106	+ .313
		" $+2S$ = .997	+.24		" $+2S$ = .987	+.197		" $+2S$ = .977	+ .184
		" $+S$ = .877	+.12		" $+S$ = .885	+.095		" $+S$ = .875	+ .082
		" $-S$ = .637	−.12		" $-S$ = .703	−.087		" $-S$ = .724	− .069
		" $-2S$ = .517	−.24		" $-2S$ = .632	−.158		" $-2S$ = .667	− .126
		" $-3S$ = .397	−.36		" $-3S$ = .561	−.229		" $-3S$ = .618	− .175
1.5	.386	Estimate $+3S$ = .746	+.36	.607	Estimate $+3S$ = .852	+.245	.621	Estimate $+3S$ = .798	+ .177
		" $+2S$ = .626	+.24		" $+2S$ = .761	+.154		" $+2S$ = .728	+ .107
		" $+S$ = .506	+.12		" $+S$ = .680	+.073		" $+S$ = .670	+ .049
		" $-S$ = .266	−.12		" $-S$ = .542	−.065		" $-S$ = .578	− .043
		" $-2S$ = .146	−.24		" $-2S$ = .484	−.123		" $-2S$ = .541	− .080
		" $-3S$ = .026	−.36		" $-3S$ = .433	−.174		" $-3S$ = .508	− .113

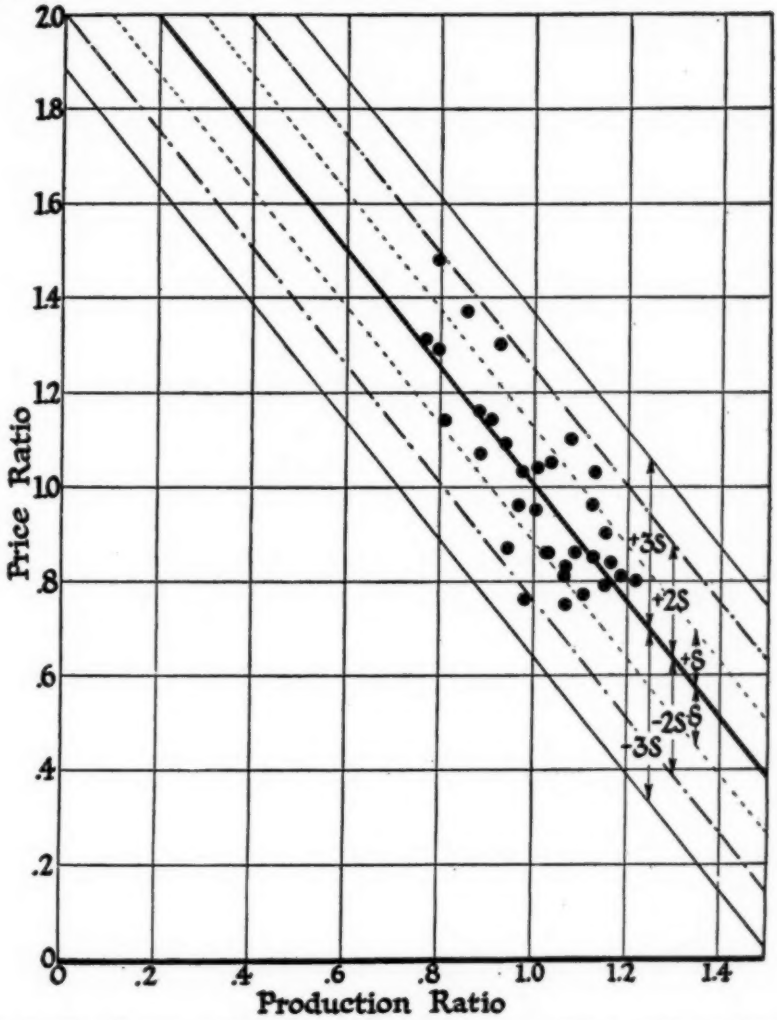


FIG. I. THE RELATION BETWEEN THE PRODUCTION AND PRICE OF OATS; ILLUSTRATING THE USE OF AN ARITHMETIC EQUATION OF REGRESSION AND ARITHMETIC ZONES OF ESTIMATE

relationship. The estimate is, in this case, the geometric mean of the actual value which exceeds it by an amount equal to the standard error of estimate (or any multiple of that quantity) and the value which falls below it by an equal amount. Figure II presents these relationships graphically. The original data are here plotted together with the graph of the equation

$$Y = .9804X^{-1.182059}.$$

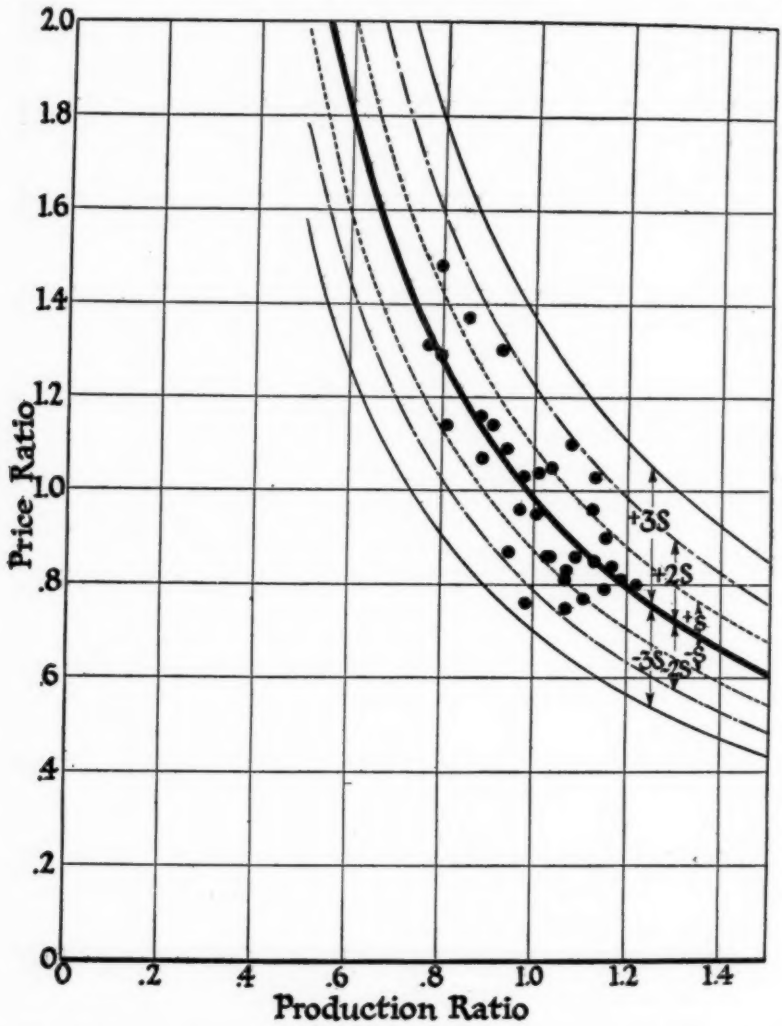


FIG. II. THE RELATION BETWEEN THE PRODUCTION AND PRICE OF OATS; ILLUSTRATING THE USE OF A LOGARITHMIC EQUATION OF REGRESSION AND GEOMETRIC ZONES OF ESTIMATE

There are shown, also, the limits of zones of estimate having widths equal, respectively, to $2S_r$, $4S_r$ and $6S_r$, centering (geometrically) at the line of relationship. A comparison of Figure I and Figure II will make clear the differences between estimates based on the assumption of an arithmetic distribution and those based on the assumption of a geometric distribution.

The points and lines shown in Figure II are plotted on a logarithmic scale in Figure III. On this scale the curve of relationship becomes straight, and the zones of estimate appear as symmetrical and of equal width throughout the range. This transformation when the data are plotted in this way reveals the fundamental simplicity of the assumptions involved in making estimates from logarithmic values.

In using the measure $S_{\frac{1}{y}}$ we carry still further the assumption that the variability about the curve is greater with high prices than with low. It shows a very limited range to errors of estimate when the estimate is low and a very wide range when the estimated price is high. A

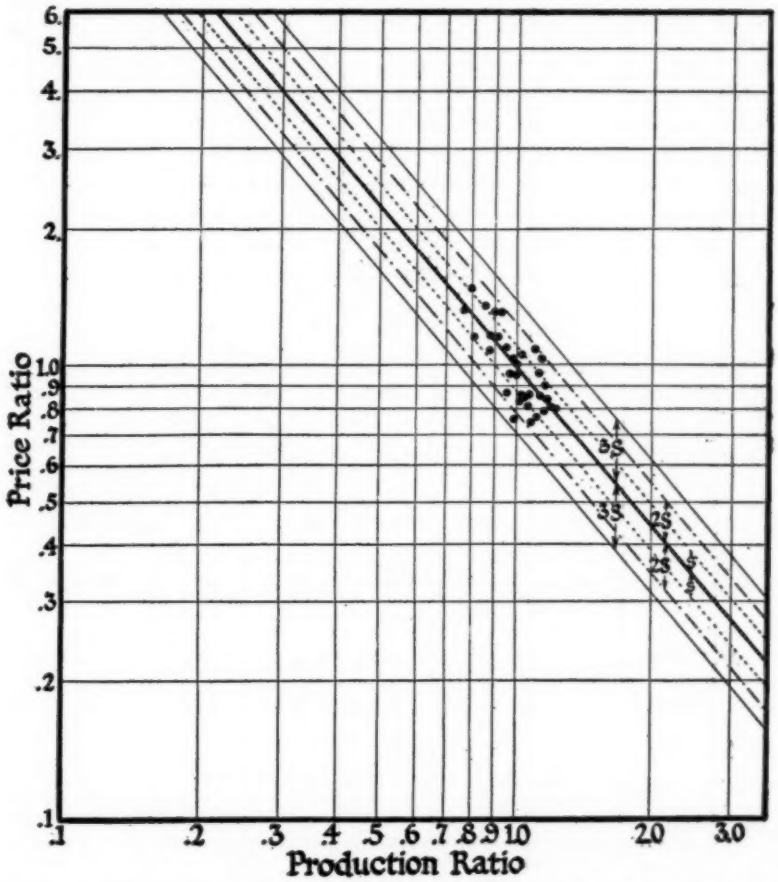


Fig. III. The Relation Between the Production and Price of Oats; Illustrating the Use of a Logarithmic Equation of Regression and Geometric Zones of Estimate (Plotted on Double Logarithmic Paper)

harmonic dispersion about the curve is assumed. The computed value, or estimate, is always the harmonic mean of the actual value which exceeds it by an amount equal to the standard error of estimate (or any multiple of that quantity) and the value which falls below it by an equal amount.

In Figure IV the curve $\frac{1}{Y} = -.1357 + 1.1643X$ is plotted, together with the original observations. Zones of estimate with widths of $2S_{\frac{1}{y}}$,

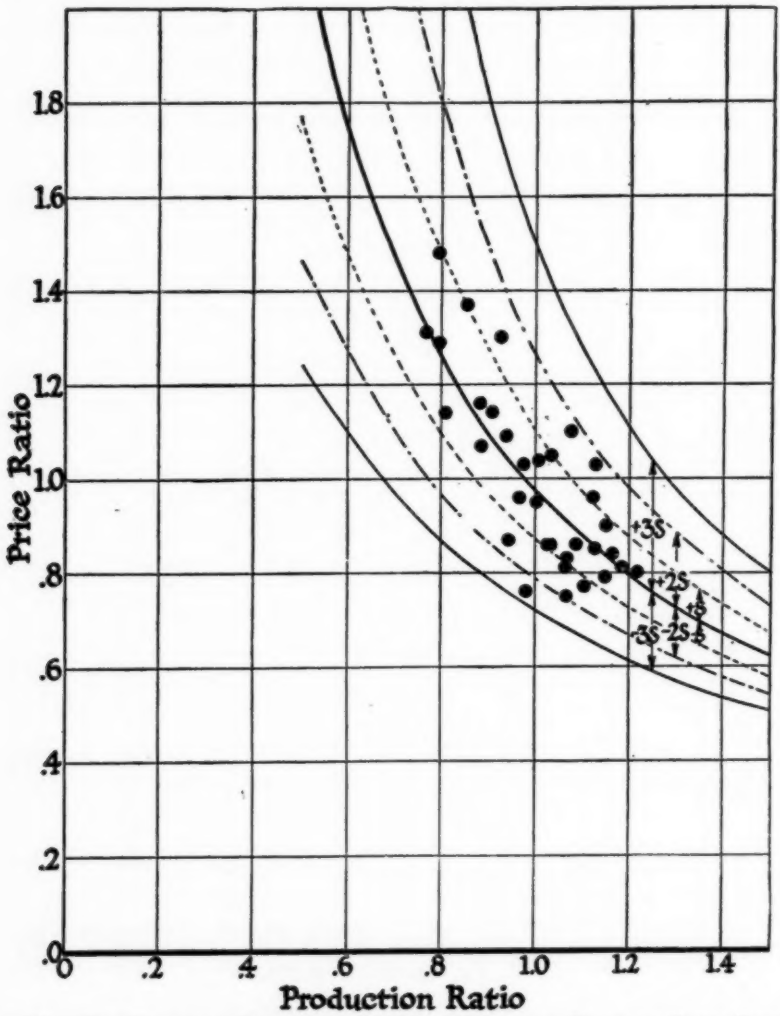


FIG. IV. THE RELATION BETWEEN THE PRODUCTION AND PRICE OF OATS; ILLUSTRATING THE USE OF AN EQUATION OF REGRESSION BASED UPON RECIPROCALS AND OF HARMONIC ZONES OF ESTIMATE

$4S_{\frac{1}{y}}$ and $6S_{\frac{1}{y}}$, centering (harmonically) at the fitted line, are shown. The differences between this figure and each of the two preceding are quite marked, particularly with respect to the zones of estimate. On the assumption of a harmonic distribution about the curve describing the relationship, the outer zone (with width equal to $6S$) marks the limits within which 99.7 per cent of all the points should fall, and the inner zone (with width equal to $2S$) marks the limits within which 68 per cent of all the points should fall. By plotting reciprocals throughout, instead of natural numbers, this apparently abnormal distribution could be reduced to the symmetrical form secured in plotting the geometric values on a logarithmic scale.

For both high and low estimates the geometric measure, $S_{\log y}$, stands between the arithmetic measure, S_y, and the harmonic measure, $S_{\frac{1}{y}}$. While the two latter have their particular functions, and are appropriate in certain cases, it is probably true that in using such methods as these in economic analysis measures of the geometric family are more generally useful than those of the other types. This means, merely, that ratios are usually more important than absolute differences. It seems reasonable, therefore, to base estimates in cases similar to the above upon an equation of the type

$$\text{Log } Y = f(X)$$

and to measure the reliability of these estimates in terms of ratios, using $S_{\log y}$ or S_r. In such cases, as we have seen, correlation is measured by an index of correlation the value of which depends upon the *ratio variability* about the curve as compared with the *ratio variability* about the geometric mean.[1]

Summary. There has been suggested in the preceding pages a general measure of correlation, to be termed the *index of correlation*, and to be represented by the symbol ρ. This measure is defined by the equation

$$\rho^2 = 1 - \frac{S^2}{\sigma^2}$$

where S^2 is the mean square deviation about a curve describing the relationship between two variable quantities and σ^2 is the mean square

[1] The reasoning in C. M. Walsh's book *The Problem of Estimation* (London, King, 1921) is peculiarly applicable to the present problem, since all such questions as those discussed above are essentially problems of estimation. Citing Galileo, in defence of the use of the geometric mean in averaging estimates, Walsh writes "And so errors must be measured by an error *which is a ratio between the estimate and the true quantity, and not a concrete quantity itself.* We cannot measure errors by so many pounds, feet or crowns, we must measure them by the proportions of the pounds, feet or crowns in the erroneous estimates to the pounds, feet or crowns in the thing estimated." (Italics mine.) (p. 12.) This argument bears out powerfully what has been said as to the use of logarithmic functions in estimating, and as to the employment of logarithmic measures of the errors of estimate.

deviation about the mean of the dependent variable. The three measures, S, σ and ρ, may be computed in terms of natural numbers, logarithms and reciprocals. In the interpretation of these values, accordingly, it is necessary to know whether they relate to arithmetic, geometric or harmonic measures. Distinctive subscripts should be employed to indicate the character of these measures in a particular instance.

In any case in which the equation to the curve employed to describe a given functional relationship may be put in the general form

$$Y = aW_1 + bW_2 + cW_3 + dW_4 + \quad . \quad . \quad .$$

the values involved in fitting by the method of least squares may be used in determining directly the sum of the squares of the residuals. This fact makes it possible to compute the values S and ρ with but little labor beyond that involved in the fitting process. Making use of these relationships, general formulae for S and ρ have been derived.

The results given above are submitted with the realization that they are in some respects tentative, and that many phases of the problem dealt with here remain to be explored. The approach seems sufficiently promising, however, to justify the presentation of this summary, in the hope that criticism and comment may bring forth whatever of significance these methods possess.

AN IMPROVED METHOD FOR MEASURING THE SEASONAL FACTOR

By W. Willford I. King

Just at present, almost every statistician is to some extent studying the business cycle. To study the cycle, it is, of course, necessary to work with concrete data.

Unfortunately for the statistician, it is rare to find figures which in their raw state actually reveal in any clear way the extent to which the fluctuations which occur in them are cyclical in origin. While numerous forces tend to obscure the cyclical movements, no other single one interferes as frequently as does the weather. The weather, then, is not only the most common of all topics of conversation, it is likewise an influence that permeates almost every phase of human activity. It is this weather influence which results in introducing into historical records of all kinds what is commonly called the seasonal factor,—a factor of such importance that its elimination is one of the first things the statistician must undertake if he really wishes to study the business cycle.

Numerous ways have been devised for taking the seasonal factor out of historical variables. One of the oldest and simplest methods is first to obtain a total of all the Januaries, a corresponding total of all the Februaries, and a similar total for each of the other months. The next step is to reduce these totals to relative numbers, the average of the totals for all the months being used as a base, and then to eliminate the seasonal by dividing each item in the original variable by the relative number for the corresponding month.

This crude method is found to be especially unsatisfactory when the data have a marked trend. If, for example, the trend is upward, it makes the January relative too low, and the December relative too high. The division of the original items by these relatives will therefore tend to produce something of a saw-tooth effect in the curve representing the resulting quotients. This difficulty can be remedied by subtracting from the item for the January just succeeding the close of the last year in the series the item for the first January in the period, and then deducting from the February total one-twelfth of the difference, from the March total two-twelfths of the difference, from the April total three-twelfths of the difference, and so on. This method approximately removes the trend from the seasonal index and thus makes it

a more trustworthy divisor for use in eliminating the seasonal factor from the original data.

The procedure just described is illustrated by the data in Table I. The average daily production of pig iron in January, 1902, was 46,380 tons, and in January, 1923, it was 104,181 tons, an increase of 57,801 tons. One-twelfth of this amount is 4,817 tons. Two-twelfths is 9,634 tons, and so on. These amounts have been subtracted from the totals before reducing the latter to relatives.

This simple method of getting rid of the seasonal factor works very well if this factor is large as compared to the influence of the cycle. If, however, the reverse is true and the seasonal fluctuations are trivial as compared to the cyclical fluctuations, the method cannot always be applied with success, for the reason that the seasonal index obtained by the process described above is frequently too inaccurate to produce satisfactory results. This inaccuracy is caused by the fact that the large cyclical movements overwhelm the small seasonal movements and hence, when the sums for all the months are obtained, their relationships to each other are dominated rather by cyclical than by seasonal movements. The resulting index in such cases is at best only a caricature of the true seasonal.

When the records for the average daily production of pig iron are plotted month by month, one sees immediately that the cyclical fluctuations are so large and the seasonal fluctuations so small as to make it not improbable that it falls in the class of variables just described, and, should such prove to be the case, it would be useless to attempt to measure the seasonal by any ultra simple methods. It may seem that in such cases the obviously wise thing to do is to ignore the seasonal altogether. In many instances, such a course may be the best policy. In others, however, it is still highly important that the seasonal be determined with as high a degree of accuracy as possible, even though its computation may require considerable effort.

For example, during the investigation into the causes of influenza, Mr. Edgar Sydenstricker of the United States Public Health Service found that it was quite essential to get rid of the seasonal influence on the rates in order to analyze with any degree of accuracy the characteristic features of epidemics of the disease. He and his assistant, Mr. Rollo Britten, therefore set about devising ways and means of accomplishing this end. After spending much time in analyzing the figures, they discovered that the seasonal factor was not constant, as has usually been assumed, but that it changed from year to year. This being the case, it was evidently inaccurate, whenever the period of time covered was considerable, to attempt its elimination through

Table I.

AVERAGE TONNAGE OF PIG IRON[a] PRODUCED DAILY IN THE UNITED STATES.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Average of twelve Months
1902	46,380	44,920	46,615	49,182	49,770	48,225	46,518	47,365	47,287	47,770	47,762	49,590	
1903	47,510	49,665	51,310	51,614	55,220	55,774	49,880	50,683	51,791	46,991	34,654	27,112	
1904	29,799	41,670	46,825	52,036	49,582	43,191	36,158	37,832	45,595	46,990	49,553	52,132	
1905	57,480	57,070	62,460	64,068	63,345	59,776	56,160	59,475	63,317	66,186	67,121	65,990	
1906	66,738	68,000	69,860	69,107	67,700	65,691	64,950	62,150	65,699	70,866	72,922	72,110	
1907	71,148	73,098	71,821	73,885	74,048	74,486	72,763	72,594	72,783	75,386	60,937	39,815	
1908	33,918	37,163	39,619	38,289	37,603	36,444	39,287	42,851	47,300	50,254	51,595	56,198	
1909	57,975	60,976	59,232	57,962	60,753	64,656	67,793	72,546	79,507	83,656	84,917	85,022	
1910	84,148	85,616	84,459	82,792	77,102	75,516	69,305	67,963	68,476	67,520	63,659	57,349	
1911	56,752	64,090	70,036	66,836	61,079	59,585	57,841	62,150	65,903	67,811	68,648	65,912	
1912	66,384	72,442	77,591	79,181	81,051	81,368	77,738	81,046	82,128	86,722	87,497	89,764	
1913	90,172	92,369	89,147	91,759	91,038	87,619	82,601	82,027	83,531	82,133	74,453	63,887	
1914	60,808	62,343	72,734	77,665	87,506	83,916	83,550	84,363	82,783	57,361	50,611	48,898	
1915	51,959	59,813	66,775	70,550	73,015	79,361	82,691	89,666	95,085	100,822	101,244	103,333	
1916	102,746	106,436	107,667	107,592	108,422	107,653	104,017	103,346	106,745	113,189	110,394	102,537	
1917	101,643	94,473	104,882	111,165	110,238	109,008	107,820	104,772	104,465	106,550	106,859	92,997	
1918	77,799	82,835	103,648	109,607	111,175	110,793	110,354	109,341	113,942	112,482	111,802	110,762	
1919	106,525	105,008	99,685	82,607	68,002	70,491	78,340	88,696	82,932	60,115	79,745	84,944	
1920	97,264	102,720	106,900	91,327	96,312	101,451	98,931	101,529	104,310	106,212	97,830	87,222	
1921	77,945	69,187	51,468	39,766	39,394	35,494	27,889	30,780	32,850	40,215	47,183	53,196	
1922	53,063	58,214	65,675	69,070	74,409	78,701	77,592	78,584	87,791	85,091	94,990	98,558	
Total	1,417,857	1,493,356	1,551,213	1,538,062	1,516,765	1,508,787	1,471,778	1,489,591	1,544,190	1,573,770	1,562,576	1,508,589	1,516,527
Deduc- tion	0	4,817	9,634	14,450	19,267	24,084	28,900	33,717	38,533	43,350	48,167	52,984	
Remain- ders	1,417,857	1,488,539	1,541,579	1,523,612	1,497,498	1,484,703	1,442,878	1,455,874	1,505,657	1,530,420	1,514,409	1,455,605	1,490,016
Ratio		.965	.999	1.016	1.021	1.005	.996	.968	.977	1.011	1.027	1.016	.977	1.000

a. Data from the Iron Age; figures cover only "coke and anthracite" pig iron.

the use of a constant seasonal index as a divisor for each year. What was needed was an index that would change from year to year as conditions changed. Manifestly, the ideal index would be one which reflected accurately every shift in weather conditions. To obtain such an index appears, however, to be entirely impossible, for one cannot tell whether an exceptional rise or fall in any one year is due to the weather or to some other cause. The best that can be hoped for is to derive an index which will show the trend of the changes occurring in the seasonal factor from year to year.

Sydenstricker and Britten were successful in working out a method which answered their needs.[1] It did not, however, appear to be entirely general in its scope. In the following pages a mode of procedure is presented which, while basically resembling their method, yet introduces certain variations in procedure which are believed to make it easier to apply and broader in its application.

Before presenting the proposed formula, the present writer wishes to state that Dr. Frederick R. Macaulay of the National Bureau of Economic Research and his assistant, Miss Gail Benjamin, some months ago worked over this same pig iron seasonal problem and by testing and weighing the devices heretofore used, and by discovering the pitfalls in the commonly followed modes of attack, made it clear to the present writer which paths ought and which paths ought not to be pursued further. One of their conclusions was that more could probably be accomplished by means of deviations from free-hand curves through the data than by means of deviations from mathematical curves or other rigid arithmetic devices.[2] It is expected that Dr. Macaulay's treatment of the whole problem of seasonals will appear in a publication of the National Bureau of Economic Research.

There has been much demand for a purely mechanical process of getting rid of the seasonal fluctuations. It appears doubtful, however, that a method of that character can be devised which is not so complex as to make the cost of applying it prohibitive. The procedure which is here presented requires for its successful use a certain degree of judgment and of familiarity with statistical curves. However, this requirement cannot be considered a matter of much real moment, for, if the statistician in charge can devote a few minutes of his time to smoothing the curves used merely by simple free-hand methods, all the rest of the work can be performed by a clerk following set instructions.

[1] So far as the present writer is aware, they were the first to devise a method of obtaining a different seasonal index for each year.

[2] The present writer had for years used the free-hand curve to eliminate any factors not wanted, and he had always believed this method to be the best possible. He was, of course, glad to find this view substantiated by the work of such an able mathematician as Dr. Macaulay.

If the statistician is willing to devote three or four hours to teaching the process to an intelligent clerical assistant, he need not even smooth the curves in the future, for the assistant can thereafter apply the method successfully without further help.

The merit of the method here set forth is that it is flexible enough to meet most tests required of such a procedure, and that the principles involved are so simple as to be easily comprehended by anyone having a good understanding of the most elementary mathematics. It gets away completely from the crude assumption that seasonal fluctuations are unchanging from year to year. The work of computation is no greater than that required by other methods which arrive only at an average seasonal, a result which, as will be shown, is distinctly inaccurate in numerous cases.

The actual procedure in obtaining the normal seasonal index for each year has been as follows:

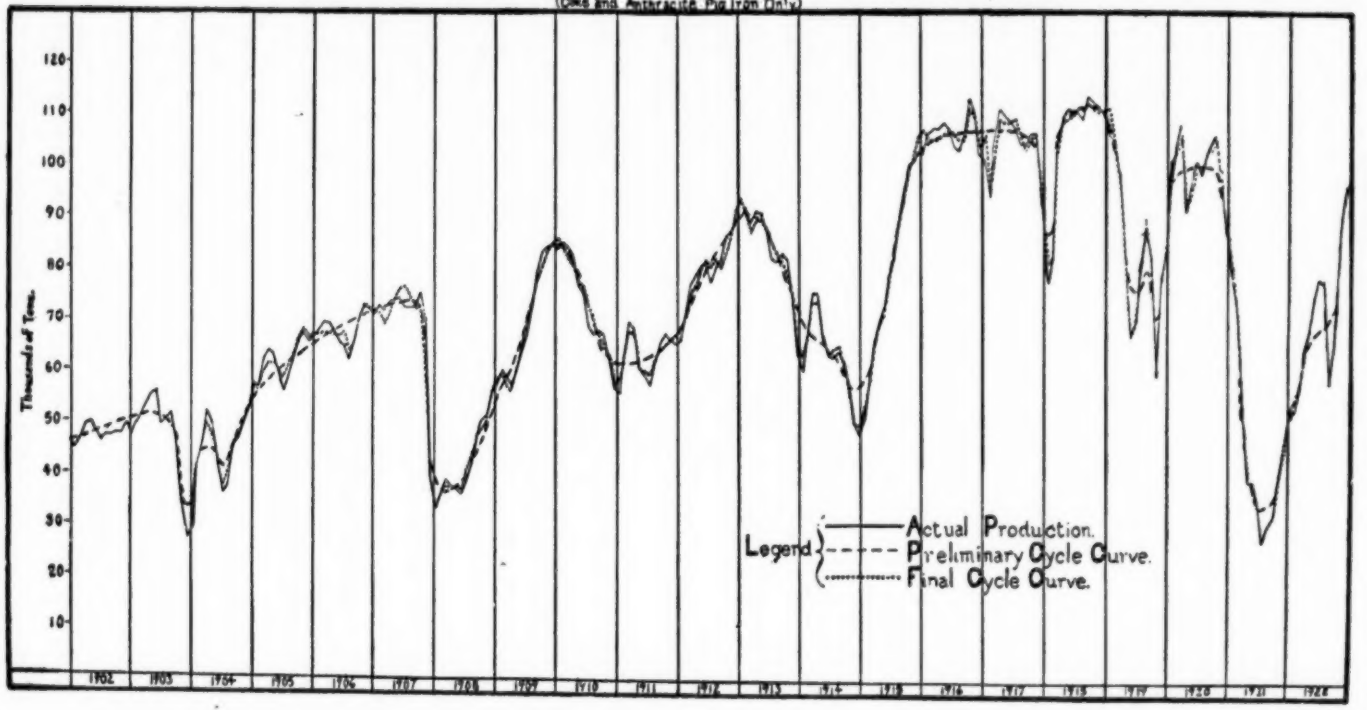
The original figures showing the average tonnage of coke and anthracite pig iron produced daily in the United States were entered in Table I. These figures were then plotted in Chart I as the solid line entitled "Actual Production." The next step after plotting the original data was to draw a free-hand curve representing what was assumed to be the course of the cycle,—in other words, what the production would have been had there been no seasonal influence at work. This curve is designated on the chart as the "preliminary cycle curve." To draw a curve of this type manifestly requires a certain degree not only of skill but likewise of discretion. For example, it is hard to say whether the peak in the early part of 1904 should be cut through or not. The same is true of the trough in the latter part of 1904. One cannot be certain how low the preliminary cycle curve should go at the end of 1914. Some light is thrown on this question by the relatives at the bottom of Table I, obtained from the totals of the respective months after adjusting for trend. These represent first approximations to the desired seasonal index, and, by observing the size of these index numbers for the various months, one is enabled to place the preliminary cycle curve with far greater accuracy than would otherwise be possible. Even with such help, however, the results brought out in this particular study show that difficulties remain. In this instance, for example, the curve as drawn allows for more seasonal influence than is actually present.

Mathematically inclined readers will doubtless view with suspicion the practice of drawing the preliminary cycle curve by the free-hand method, and will inquire why a mathematical process is not used. The answer is that no simple mathematical process seems to be even mod-

Chart I.

ACTUAL AND CYCLICAL MOVEMENTS IN THE MONTHLY AVERAGE OF THE DAILY PRODUCTION OF PIG IRON IN THE UNITED STATES.

(Coke and Anthracite Pig Iron Only)



erately well adapted to this end,[1] and the present writer doubts that the most intricate one that can be devised will serve the intended purpose as well as a free-hand curve placed by an intelligent person moderately familiar with graphic statistics. It is realized that this is a point upon which opinions may differ,—but the only way to prove erroneous the view that the free-hand curve is superior, is to produce the mathematical curve that is obviously a better fit.

The next step taken toward obtaining the seasonal index was to read from the free-hand curve the figures representing for each month the tentative estimate of the location of the cycle. These figures are entered in Table II. The items in Table I were then divided by those for corresponding dates in Table II. The quotients thus obtained (which appear in Table III), represent an approximation to the seasonal index numbers for the various months.

TABLE II

READINGS FROM FREE-HAND CURVE SHOWING TENTATIVE ESTIMATE OF DAILY PIG IRON PRODUCTION AFTER SEASONAL FLUCTUATIONS HAVE BEEN ELIMINATED

(Thousands of tons)

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1902.....	46.4	46.8	47.0	47.4	47.8	48.1	48.5	48.8	49.2	49.6	50.0	50.3
1903.....	50.6	50.8	51.0	51.4	51.5	51.3	50.9	50.1	49.0	47.5	37.6	33.4
1904.....	35.0	40.4	44.6	45.1	44.9	44.3	41.6	42.2	46.0	48.6	50.9	53.1
1905.....	54.9	56.0	57.3	58.4	59.4	60.2	61.0	61.9	62.8	63.6	64.1	64.8
1906.....	65.5	66.0	66.9	67.5	68.0	68.6	69.1	69.8	70.2	70.9	71.3	71.8
1907.....	72.1	72.5	72.9	73.0	73.3	73.5	73.7	73.8	73.6	72.0	67.0	43.0
1908.....	39.5	37.5	37.0	37.3	38.1	39.0	40.0	42.0	44.2	46.4	49.4	52.0
1909.....	55.3	57.2	59.3	61.2	64.0	67.1	70.3	74.1	77.9	80.9	83.9	84.9
1910.....	85.0	84.8	83.4	81.4	79.3	76.9	73.2	70.2	66.7	64.4	62.9	62.4
1911.....	62.0	62.0	62.1	62.2	62.6	62.9	63.3	64.0	64.5	65.6	66.5	67.8
1912.....	69.1	71.0	73.2	76.0	77.7	79.7	81.7	83.3	85.0	86.7	87.7	89.5
1913.....	90.5	91.0	90.9	90.2	89.5	88.0	85.9	83.0	80.0	77.2	75.3	73.0
1914.....	70.3	68.6	67.5	66.5	65.8	64.9	63.9	62.2	60.2	58.1	57.2	57.7
1915.....	59.1	61.6	65.0	68.9	73.0	80.1	85.0	89.5	95.0	98.1	101.4	103.0
1916.....	104.0	104.9	105.6	105.9	106.2	106.4	106.6	106.8	106.9	107.0	107.1	107.2
1917.....	107.2	107.3	107.4	107.4	107.3	107.1	106.9	106.8	106.3	105.5	104.0	92.0
1918.....	87.2	88.9	105.0	108.6	109.9	110.5	111.3	111.6	111.8	111.7	111.2	110.4
1919.....	107.8	103.2	96.0	82.8	76.8	76.0	76.7	80.2	76.0	71.0	78.3	87.0
1920.....	96.1	98.9	99.9	100.1	100.3	100.3	100.2	99.9	99.3	96.9	92.0	87.0
1921.....	81.4	69.0	47.0	41.2	37.0	34.7	34.4	34.9	36.3	41.0	47.1	51.4
1922.....	55.4	58.0	64.7	66.4	67.9	68.8	69.6	71.0	72.9	86.0	93.0	98.7

Owing to the impossibility of placing the preliminary cycle curve with any high degree of accuracy, and on account also of the numerous "accidental" forces playing a part, the figures in Table III show a very considerable degree of irregularity. As stated above, since it is not feasible to determine the actual seasonal, the object sought is to obtain the normal seasonal for each date. The first attack on the problem

[1] This conclusion is based upon the previously referred-to studies made by Dr. F. R. Macaulay.

TABLE III

RATIOS OF ACTUAL TO TENTATIVE CYCLICAL DAILY PRODUCTION OF PIG IRON

(Items in Table I divided by corresponding items in Table II)

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Av. of mos.
1902	1.000	.960	.992	1.038	1.041	1.002	.959	.971	.961	.963	.955	.986	.986
1903	.939	.977	1.006	1.043	1.072	1.087	.979	1.012	1.057	.968	.922	.818	.990
1904	.852	1.032	1.050	1.154	1.104	.975	.869	.897	.991	.966	.973	.982	.987
1905	1.047	1.019	1.091	1.087	1.067	.993	.921	.961	1.008	1.046	1.046	1.018	1.025
1906	1.019	1.030	1.044	1.024	.946	.961	.940	.890	.935	.999	1.023	1.004	.985
1907	.987	1.007	.985	1.012	1.013	1.013	.987	.984	.988	1.047	.909	.926	.988
1908	.859	.991	1.071	1.027	.987	.934	.982	1.020	1.070	1.089	1.044	1.080	1.013
1909	1.048	1.066	.999	.947	.949	.963	.964	.979	1.021	1.036	1.012	1.013	1.000
1910	.990	1.010	1.012	1.017	.972	.982	.947	.968	1.026	1.048	1.012	.919	.992
1911	.915	1.034	1.128	1.107	.976	.947	.914	.971	1.022	1.034	1.002	.972	1.003
1912	.961	1.020	1.060	1.042	1.043	1.021	.952	.973	.966	1.000	1.000	1.003	.998
1913	.996	1.015	.981	1.017	1.017	.996	.962	.989	1.044	1.064	.989	.876	.995
1914	.865	.983	1.122	1.138	1.026	.985	.988	1.035	1.043	.988	.885	.847	.992
1915	.874	.971	1.024	1.023	1.000	.991	.973	1.002	1.001	1.028	.998	1.003	.982
1916	.989	1.015	1.020	1.016	1.021	1.005	.976	.968	.999	1.058	1.031	.957	1.004
1917	.948	.881	.977	1.035	1.027	1.018	1.009	.981	.983	1.010	1.028	1.017	.993
1918	.892	.932	.987	1.009	1.012	1.003	.991	.980	1.019	1.007	1.006	1.003	.987
1919	.988	1.018	1.039	.998	.885	.927	1.021	1.103	1.091	.846	1.019	.976	.992
1920	1.012	1.039	1.091	.912	.960	1.012	.991	1.016	1.050	1.096	1.063	1.003	1.020
1921	.958	1.003	1.095	.965	1.063	1.023	.811	.882	.905	.981	1.002	1.035	.977
1922	.938	1.004	1.015	1.040	1.096	1.144	1.115	.825	.930	.989	1.021	1.009	1.012

was made by calculating a moving average for all the Januaries, a similar one for all the Februaries, and so one. This method proved unsatisfactory because a few of the irregularities were so large as to make it almost impossible to obtain smooth curves. After several attempts, it was found that the best way of attaining the desired end was to apply to all of the Januaries a nine-period moving median.[1] This item was plotted at the middle year of the period: The same process was then applied to the Februaries,—and so on for the rest of the months. These curves were then plotted in a manner similar to that shown in Charts 2a and 2b. It was still found that certain chance irregularities remained and these were smoothed out. Manifestly, if a seasonal index having unity as a base is accurate, the sum of the index numbers for the twelve months should equal exactly twelve. The various curves were therefore adjusted slightly up or down until the totals were brought to this figure. The process of adjusting these monthly curves is the second place in which a certain degree of discretion is required. In making the adjustment, it is necessary to consider the fact that presumably the fluctuations in adjacent months will resemble each other to some degree and that a month intervening between two others is likely to partake, to a considerable extent, of the

[1] The present writer believes this to be an entirely new use of the moving median—a most usefu[l] device, the value of which is apparently yet to be recognized by statisticians.

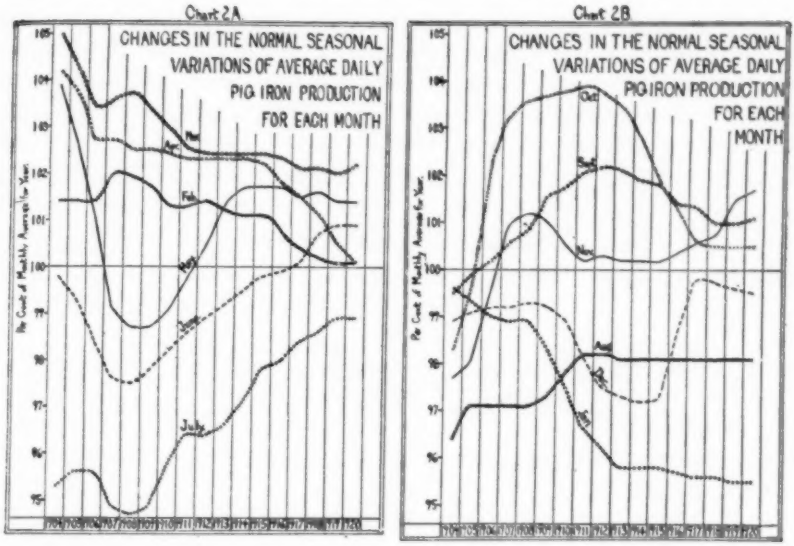


Chart 2A
CHANGES IN THE NORMAL SEASONAL VARIATIONS OF AVERAGE DAILY PIG IRON PRODUCTION FOR EACH MONTH
Chart 2B
CHANGES IN THE NORMAL SEASONAL VARIATIONS OF AVERAGE DAILY PIG IRON PRODUCTION FOR EACH MONTH

characteristics of the months on either side. As a matter of fact, in most of the years the adjustments required were only minor in extent and relatively little skill was necessary in order to secure fairly smooth curves and at the same time make the sum of the readings for the year total twelve. From the curves presented in Charts 2a and 2b, one can

TABLE IV

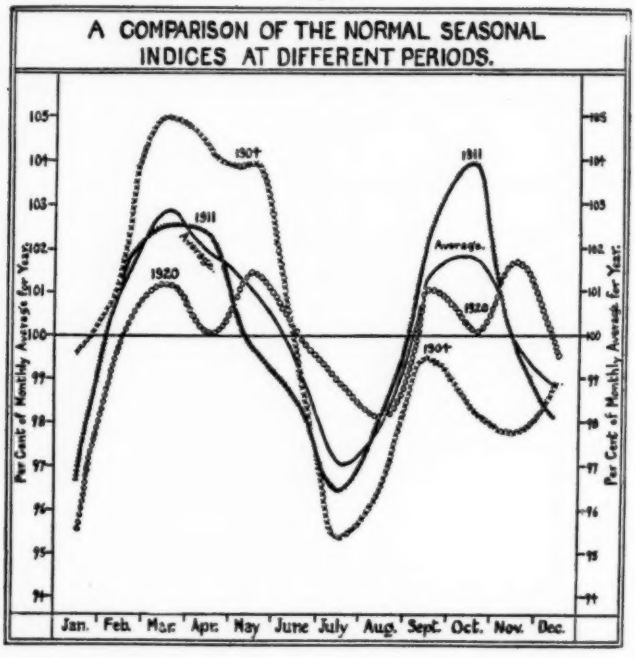
NORMAL SEASONAL FLUCTUATIONS IN THE PRODUCTION OF PIG IRON

(Figures read from curves in Charts 2A and 2B; expressed as per cents of the monthly average for the year)

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Totals
1904.....	.996	1.014	1.050	1.042	1.039	.998	.953	.964	.995	.983	.977	.989	12.000
1905.....	.993	1.014	1.042	1.037	1.026	.993	.956	.971	.999	.997	.981	.991	12.000
1906.....	.990	1.014	1.034	1.027	1.010	.984	.956	.971	1.002	1.024	.996	.992	12.000
1907.....	.989	1.020	1.036	1.027	.991	.976	.949	.971	1.006	1.033	1.010	.992	12.000
1908.....	.989	1.019	1.037	1.025	.987	.975	.947	.971	1.009	1.036	1.012	.993	12.000
1909.....	.983	1.017	1.033	1.025	.987	.978	.949	.973	1.016	1.037	1.010	.992	12.000
1910.....	.975	1.013	1.029	1.024	.991	.982	.958	.978	1.018	1.038	1.005	.989	12.000
1911.....	.966	1.013	1.025	1.023	.998	.986	.964	.982	1.021	1.039	1.002	.981	12.000
1912.....	.962	1.014	1.024	1.023	1.005	.989	.964	.982	1.022	1.037	1.003	.975	12.000
1913.....	.958	1.012	1.024	1.023	1.014	.992	.966	.981	1.021	1.034	1.002	.973	12.000
1914.....	.958	1.011	1.024	1.023	1.017	.995	.971	.981	1.019	1.027	1.002	.972	12.000
1915.....	.958	1.011	1.024	1.022	1.017	.998	.978	.981	1.018	1.018	1.002	.973	12.000
1916.....	.957	1.007	1.023	1.018	1.017	.999	.980	.981	1.014	1.012	1.004	.988	12.000
1917.....	.956	1.004	1.021	1.015	1.015	1.001	.984	.981	1.013	1.006	1.006	.998	12.000
1918.....	.956	1.002	1.021	1.011	1.016	1.007	.986	.981	1.010	1.005	1.008	.997	12.000
1919.....	.955	1.001	1.020	1.005	1.014	1.009	.989	.981	1.010	1.005	1.015	.996	12.000
1920.....	.955	1.001	1.022	1.001	1.014	1.009	.989	.981	1.011	1.005	1.017	.995	12.000
1921.....	.955	1.001	1.022	1.000	1.014	1.009	.989	.981	1.011	1.005	1.019	.994	12.000
1922.....	.955	1.001	1.022	.999	1.014	1.009	.989	.981	1.011	1.005	1.021	.993	12.000
Average.....	.969	1.010	1.028	1.019	1.010	.994	.970	.978	1.012	1.018	1.005	.988	1.000

read the approximately correct normal seasonal for any month in the period from 1904 to 1920, inclusive. The use of the moving median, like that of the moving average, has the weakness that it does not cover the extreme items. Therefore, to make an estimate for the most recent years, it is necessary to produce the curves from their termini. While this cannot be done with a high degree of accuracy, it is believed that an examination of the charts will lead to the conclusion that the error arising from this difficulty is certainly very much less than the error present if one used the average seasonal index for all years.

Chart 3.



Charts 2a and 2b show, for example, that, during the seventeen years under discussion, the seasonal index for a given month was sometimes well above normal and at other times well below, and that almost every one of the months shows a decidedly variable trend. In Chart 3, the seasonals for three different years are presented and compared with the average seasonal for all years. Obviously, there is a wide difference between the seasonal for 1904 and that for 1920, and neither of them resembles much the average seasonal for the entire period. Under such circumstances, the use of the average seasonal for all years certainly has little to commend it.

TABLE V

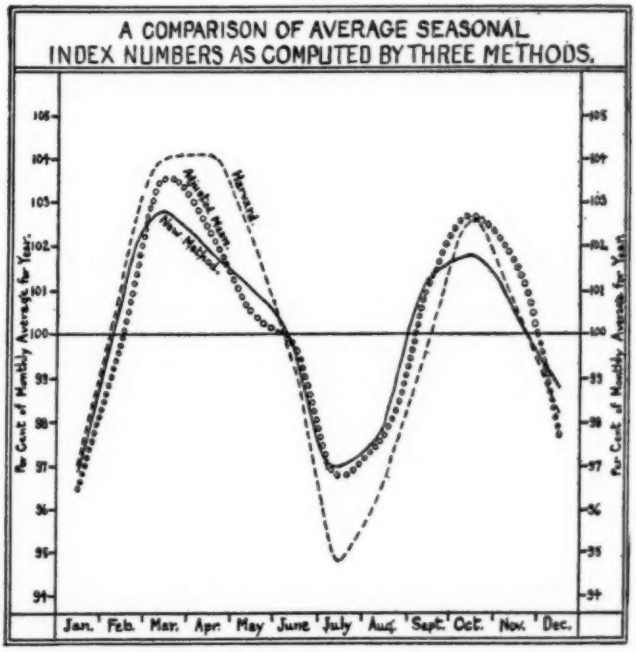
A COMPARISON OF AVERAGE SEASONAL INDICES FOR PIG IRON PRODUCTION AS SHOWN BY THREE METHODS

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Adjusted average *	.965	.999	1.036	1.023	1.005	.996	.968	.977	1.011	1.027	1.016	.977
New method †	.969	1.010	1.028	1.019	1.010	.994	.970	.978	1.012	1.018	1.005	.988
Harvard ‡	.970	1.012	1.040	1.041	1.021	.991	.948	.966	.997	1.026	1.006	.982

* See Table I.
† See Table IV.
‡ See Warren M. Persons, *The Review of Economic Statistics*, January 1919, p. 31. The figures have been adjusted for the varying lengths of the respective months in order to make them comparable with the indices calculated by the other methods which are based upon the average daily production of pig iron.

Chart 4 compares the average seasonal for pig iron production as computed by the Harvard University Committee on Economic Research in their "Review of Economic Statistics" for January, 1919, p. 30,[1] with the average seasonal computed by the old method and shown at the bottom of Table I and also with that obtained by averaging the independent seasonals for the various years as recorded in Table IV.

Chart 4.


A COMPARISON OF AVERAGE SEASONAL INDEX NUMBERS AS COMPUTED BY THREE METHODS.
Per Cent of Monthly Average for Year.
Harvard
Adopted Means
New Method
Jan. Feb. Mar. Apr. May June July Aug. Sept. Oct. Nov. Dec.

[1] The Harvard figures have been adjusted for the varying lengths of the respective months in order to make them comparable with the indices here given which are computed on the basis of average daily production of pig iron.

It appears that the averages obtained by the three methods do not differ greatly. The figures in Table V show that, in toto, the simple adjusted mean as recorded in Table I diverges from the average of Table IV by little more than half as much as does the Harvard index. It appears, therefore, that when data for a considerable period of years are available, it is superfluous, even when dealing with a difficult case like that of pig iron, to use any but the simplest procedure, if one desires only to obtain an *average* seasonal index. The game is not worth the candle.

However, the reader should not be led by this discussion of the relative merits of averages obtained in different ways to believe that even the one derived by the approved method will produce satisfactory results when used as a divisor for eliminating the seasonal in all the different years. Such a use manifestly involves the tacit assumption that the seasonal does not change materially with the passage of time. The experience of the present writer indicates that, very frequently, this assumption is decidedly at variance with the truth. If such is the case, evidently one can be assured of a reasonable degree of accuracy in the results only when the divisor used to remove the seasonal influence varies from year to year in accordance with the facts. To use an average seasonal without investigation is, then, merely a time-saving device which is permissible only in those cases in which rough esti-

TABLE VI

CYCLICAL MOVEMENTS IN THE AVERAGE DAILY PRODUCTION OF PIG IRON IN THE UNITED STATES

(Results after the elimination of the normal seasonal)

(Thousands of tons)

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Av.
1904	29.9	41.1	44.6	49.9	47.7	43.3	37.9	39.2	45.8	47.8	50.7	52.7	44.2
1905	57.9	56.3	59.9	61.8	61.7	60.2	58.8	61.3	63.4	66.4	68.4	66.6	61.9
1906	67.4	67.1	67.6	67.3	67.0	67.0	67.9	64.0	65.6	69.2	73.2	72.7	68.0
1907	71.9	71.6	69.3	71.9	74.7	76.3	76.7	74.8	72.3	73.0	60.3	40.1	69.4
1908	34.3	36.5	38.2	37.4	38.1	37.4	41.5	44.1	46.9	48.8	51.0	56.5	42.6
1909	59.0	59.9	57.3	56.5	61.6	66.1	71.4	74.6	78.3	80.9	84.1	85.7	69.6
1910	86.3	84.5	82.1	80.9	77.8	76.9	72.3	69.5	67.3	65.0	63.3	58.0	73.7
1911	58.8	63.3	68.3	67.3	61.2	60.4	60.0	63.3	64.5	65.3	66.5	67.2	63.8
1912	69.0	71.4	75.8	77.4	80.6	82.3	80.6	82.5	80.4	83.6	87.4	92.1	80.3
1913	94.1	91.3	87.1	89.7	89.8	88.3	85.5	83.6	81.8	79.4	74.3	65.8	84.2
1914	63.5	66.7	74.0	74.0	66.4	64.2	65.0	65.6	61.6	55.9	50.5	50.3	63.3
1915	53.9	59.2	65.0	69.0	71.8	79.5	84.6	91.4	93.4	99.0	101.0	106.2	81.2
1916	107.4	105.7	105.2	105.7	106.6	107.2	106.2	105.4	105.3	111.8	109.9	103.8	106.7
1917	106.3	94.1	102.7	109.5	108.6	108.9	109.6	106.8	103.1	105.9	106.2	93.2	104.6
1918	81.4	82.7	101.5	108.4	109.4	110.0	111.9	111.5	112.8	111.9	110.9	111.1	105.3
1919	111.5	105.0	97.7	82.2	67.1	69.9	79.2	90.2	82.1	59.8	78.6	85.3	84.0
1920	101.9	102.6	106.6	91.2	95.0	100.5	100.0	103.5	103.2	105.7	96.2	87.7	99.5
1921	81.6	69.1	50.4	39.8	38.9	35.2	28.2	31.4	32.5	40.0	46.3	53.5	45.6
1922	55.6	58.2	64.3	69.1	73.4	78.0	78.5	59.7	67.1	84.7	93.0	100.3	73.4

mates are sufficient. To ascertain the facts often necessitates the computation of the varying seasonal!

The last step in the process of obtaining a record of the cyclical production of pig iron has been to divide each of the original items in Table I by the seasonal index shown in Table IV for the corresponding month of the same year. The quotients, which represent the final approximation to the cyclical movements in pig iron production are entered in Table VI and shown in the dotted line in Chart I. A glance at this chart shows that, in the preliminary guess, the extent of the seasonal fluctuations was materially overestimated, and it was this overestimate that made the readings in Table III so irregular. However, the guessed cycle curve was much closer to the true cycle curve than a moving average line or any other simple mathematical curve would be. The greater this proximity, the less, of course, is the error in the final results. When one observes how close the final "cycle" or residual curve is to the original graph of pig iron production, one realizes more fully than ever the obstacles involved in arriving at a reasonably accurate series of seasonal index numbers and the necessity of using for the purpose as accurate a method as possible.

The results also demonstrate the desirability of measuring the extent of the seasonal fluctuations even when they are as small as they are in the case of pig iron production. The location of the preliminary cycle curve shows that when it was drawn many of the smaller cyclical or fortuitous movements were assumed to be merely seasonal manifestations. The final "cycle" curve therefore includes many minor waves which do not appear in the preliminary estimate. It is evident, then, that the more accurate placing of the curve clarifies our ideas, not only concerning the size of the seasonal, but also in regard to the nature of the residual curve.

To sum up, it is believed that the new method here presented not only has an excuse for existence but will prove itself superior to any heretofore published because:

1. It is easy to understand.
2. It can be computed with a reasonable degree of effort.
3. It has the outstanding merit of giving a separate seasonal index for each year during the period.

While other methods may be quite satisfactory for cases in which the seasonal happens to remain unchanged and in which it also dominates the cyclical influence, it appears that some such method as the one just set forth is far superior for the solution of problems of the more difficult type, as, for example, the derivation of the seasonal index for pig iron production.

IS THE FAMILY OF FIVE TYPICAL?

By Paul H. Douglas, *University of Chicago*

Modern budgetary studies almost universally adopt as their norm the family of five, composed of husband, wife, and three children under 14. The wages of adult males are customarily measured against the cost of maintaining such a family and are adjudged deficient or adequate according to their deviations from it. The representatives of labor in arbitration proceedings commonly submit the cost of living for such a family as furnishing the minimum upon which wages should be based and arbitrators generally take this factor at least into consideration in making their awards.

It is important therefore to scrutinize carefully the validity of the so-called standard family of five. This size of family has apparently been used for two main reasons: first, it was believed to be approximately the average sized family and, secondly, a family was supposed on the average to need at least three children in order to prevent a decrease in the population. It is desirable however to check these assumptions by statistical facts to determine how typical the family of five actually is and it is that which this article attempts to do.

1. THE SIZE OF FAMILIES IN GREAT BRITAIN

There is more complete data on the relative size of families in England than for any other country and for that reason, we shall begin there. The first thorough analysis of the number of children dependent upon the workers for support was made by Bowley and Burnett-Hurst in *Livelihood and Poverty*, based upon a survey of four English cities.[1] Approximately one out of every twenty working-class households in these localities was visited: the composition of approximately 2,000 families with male wage-earners was then analyzed with the following results.[2]

Number of dependent children all ages per family	Percentage of families with male wage-earner having given number of dependent children
0	30.2
1	21.9
2	18.8
3	13.0
4	8.2
5	4.0
6 (or more)	3.9

[1] Northampton: Warrington, Stanley and Reading.

[2] Computed from data in Bowley and Burnett-Hurst, *Livelihood and Poverty*, p. 30.

Since the household was made the unit and not the worker, the inclusion of single workers without dependents would have appreciably increased those with no dependent children. Even as it was, however, the family of three children was found to be characteristic of only 13 per cent of the households while 70.9 per cent had fewer than this number and 16.1 per cent had more.

Bowley followed up this study in 1921 by sampling one householder's schedule in the 1911 census out of fifty in seven additional cities. All middle-class families were eliminated and data on working-class households alone were compiled.[1] An elaborate method of classification was used: and the following table summarizes his results.[2]

TABLE I

RELATIVE NUMBER OF DEPENDENTS IN HOUSEHOLDS OF SEVEN ENGLISH CITIES, 1911

Number of dependents* (*i. e.*, non-earners)		Percentage of households (excluding broken families and those where no male was employed)	
Below 14	Above 14	Skilled	Unskilled
0	0	8	9
0	1	26	26
0	2 (or more)	5	3
1	0	2	2
1	1	17	16
1	2 (or more)	2	2
2	0	1	2
2	1	14	13
2	2 (or more)	2	1
3	0	0	1
2	1	9	9
3	2 (or more)	2	1
4	1	6	6
4	2 (or more)	2	1
5	1	4	5
5	2	1	2

*By dependents is meant a non-wage earner; *i. e.*, a wife at work is not classed as a dependent. This prevents the table itself from showing how many should be dependents. Less than 2 per cent of the wage earners, however, were under 14.

Bowley went on to point out that assuming "the only person over 14 properly dependent on the man is his wife and also that the father is solely responsible for the support of all children under 14, and that a wage is fixed adequate for a man, wife and three children," then, after eliminating the broken families and those where no man was working, "the wage would be exactly adequate in only 11 per cent of the households; it would be inadequate in 13 per cent of the cases and would allow a margin in 76 per cent."[3]

Bowley found that the number of workmen, including both married

[1] A. L. Bowley, "Earners and Dependents in English Towns," *Economica*, May, 1921 (Vol. 1).
[2] Computed from data given by Bowley, *ibid.*, pp. 107–8.
[3] *Ibid.*, pp. 108–09.

and single, out of every thousand, classified according to the number of children under 14 years dependent upon them for support was as follows.[1]

Number of dependent children under 14 years	Number of workmen out of 1,000 with given number of dependents
0	531
1	162
2	125
3	84
4	55
5	28
6	10
7	4
8 or more	1

This showed that only 8.4 per cent had three such children dependent upon them, while 81.8 per cent had less than this number and 9.8 per cent had more. Over 50 per cent had no children at all.

As a result of the general interest as to the relative number of dependents in the families, the British census in 1921 added a question calling for the number of children under 16 years of each married man or woman. When this was added to the statistics gathered concerning the number of single men and women in comparison with the number married, widowed or divorced, there was made possible for the first time a picture of the family composition of the households of millions of workers. The British census returns are being published separately by counties and the combined results for the country as a whole have not yet been issued. The returns for the first fourteen counties covering approximately 6.8 million gainfully occupied males of 20 years and over which are given below are, however, without doubt representative of conditions in general.

The following table shows the number of single, married, widowed and divorced adult males who were gainfully employed and the proportion which each formed of the total.[2]

Slightly over 26 per cent of all the adult males who are gainfully occupied are therefore unmarried. Only an insignificant proportion of these have children to support although more doubtless, have adults who are dependent upon them. Few, however, have the equivalent of a family of five to care for and it seems probable that the majority have no one completely dependent upon them. A wage sufficient to

[1] A. L. Bowley, "Earners and Dependents in English Towns," *Economica*, May, 1921 (Vol. 1), p. 111.

[2] The number of occupied males who were either married, or widowed and divorced is given directly for each county. The total number of males of 12 years and over is also given. By subtracting the number of males of the years 12–19 inclusive less the number of married juveniles from this figure, one secures the number of adult unmarried males of 20 years and over who are gainfully employed.

TABLE II

MARITAL CONDITION OF 6.8 MILLION ADULT OCCUPIED MALES (20 YEARS AND OVER) IN 14 ENGLISH AND WELSH COUNTIES—1921

County	Number of occupied males 20 years and over according to marital condition (in 000's)				Percentage of occupied males 20 years and over according to marital condition		
	Single	Married	Widowed or divorced	Total	Single	Married	Widowed or divorced
1. London	324	821	55	1,200	27	68	5
2. Middlesex	69	240	12	321	21	75	4
3. Durham	127	272	19	418	30	65	5
4. Essex	100	277	16	393	25	71	4
5. Glamorgan	107	234	17	358	30	65	5
6. Hertford	22	62	3	87	26	70	4
7. Kent	80	211	13	304	26	70	4
8. Lancaster	367	928	67	1,362	27	68	5
9. Monmouth	38	85	6	139	29	66	5
10. Stafford	93	256	17	366	25	70	5
11. Surrey	61	172	9	242	25	70	5
12. Warwick	94	272	17	383	24	71	5
13. Worcester	28	76	5	109	25	70	5
14. Yorkshire	305	817	55	1,177	26	69	5
	1,815	4,723	311	6,849	26.5	69.0	4.5

maintain a family of five would therefore be vastly in excess of the present needs of this group, comprising as it does over one-quarter of all the adults who are at work.

The married and widowed workers in each county are also classified according to the number of children under 16 that they have, as is shown in Table III. Due to the newness and complexity of the question calling for the number of children, it is probable that a consider-

TABLE III

RELATIVE NUMBER OF CHILDREN UNDER 16 YEARS PER FAMILY IN ENGLAND AND WALES—1921

County	Percentage of married men or widowers with varying number of living children under sixteen										
	No children or no statement	1	2	3	4	5	6	7	8	9	10 or more
1. London	43.8	22.6	15.0	8.7	5.0	2.8	1.4	.5	.2	0	0
2. Middlesex	42.7	24.0	16.2	8.6	4.4	2.3	1.1	.5	.2	0	0
3. Durham	34.6	22.9	17.0	11.4	7.1	4.1	2.0	.7	.2	0	0
4. Essex	42.2	23.0	15.7	9.0	5.1	2.8	1.4	.5	.2	.1	0
5. Glamorgan	35.8	23.1	16.6	10.7	6.6	3.9	2.0	.9	.3	.1	0
6. Hertford	45.7	22.7	15.0	8.0	4.5	2.3	1.1	.5	.2	0	0
7. Kent	46.0	22.5	14.7	8.1	4.4	2.4	1.2	.5	.2	0	0
8. Lancaster	42.5	23.9	15.5	8.8	4.9	2.6	1.2	.5	.1	0	0
9. Monmouth	34.4	23.3	17.0	11.1	6.8	4.0	2.1	.9	.3	.1	0
10. Stafford	38.0	23.1	16.2	10.1	6.2	3.6	1.8	.7	.2	.1	0
11. Surrey	46.5	23.0	15.1	7.8	4.0	2.1	1.0	.4	.1	0	0
12. Warwick	40.9	23.9	16.1	9.1	5.1	2.8	1.3	.5	.2	.1	0
13. Worcester	43.3	23.0	15.1	8.9	5.0	2.7	1.3	.5	.2	0	0
14. Yorkshire	43.1	23.7	15.3	8.7	4.8	2.6	1.2	.5	.1	0	0
Total	41.9	23.3	15.6	9.1	5.2	2.9	1.4	.5	.2	0	0

able proportion of parents did not make full reports and that an appreciable number of those who actually had children did not make any such statement. The British census, however, believes that a "reasonable degree of confidence" may be placed in the results for the married people, "accounting as they do for about 90 per cent of the children."[1]

The fact that they roughly check with Bowley's earlier study of ten thousand workers gives them a further degree of corroboration.

Forty-two per cent of the married workers would seem, therefore, to have had no dependent children at the time that the inquiry was made;[2] 23 per cent had only one child and 16 per cent had two, 81 per cent in all, therefore, had less than the number commonly set as the standard. Only 9 per cent of those that were married had the standard of three children while 10 per cent had more than this number.

Since the married and widowed workers formed only 73.5 per cent of all the adult males, it will be seen that the so-called standard family was in reality characteristic of only 7.6 per cent of all the workers. The average number per family of children under sixteen years in the fourteen counties was indeed only 1.31.

2. FRANCE AND BELGIUM

There have been two interesting investigations into the size of French families. The first was that conducted in 1919 by a group of Parisian manufacturers into the family status of 99,600 workmen who were almost exclusively employed in the metallurgical and metal industries of the capital. They found that 36.9 per cent of the total were either unmarried or widowed or divorced without having dependent children. On the other hand 63.1 per cent were married.[3]

These married workmen in turn had the following number of children dependent upon them:

Number of children	Per cent of married workmen
0	38.4
1	33.6
2	17.5
3	6.4
4	2.5
5 (or more)	1.6
Total	100.0

[1] See report on *County of Middlesex*, Census of England and Wales, 1921 (1923) p. XXXI. Similar statements are made in the case of most of the other counties.

[2] This is probably something of an overestimate since it includes those who made no statement but yet had dependent children.

[3] See Pierre Richemond, Allocations Pour Charges de Famille et Caisses de Compensation, *Revue D'Economie Politique*, Vol. XXIV, pp. 590–91.

The family of three children thus seemed to be characteristic of only 6.4 per cent of the 63 per cent of married workers or only 4 per cent of the total number of workers, while less than this number had more than three children dependent upon them for support.

The second investigation was that made in 1923 by Dr. Perret of Lyons for the Comité des Allocations Familiales and which covered several hundred thousand employees. Dr. Perret found that only 24.7 per cent of the workers were fathers of children.[1] Since a large number of women and juveniles, however, were included as workers, this figure should not be greatly relied upon. The average number of children in families where there were any at all was 1.66 and the proportion which had the varying number of children was as follows:

Number of children in family	Per cent of families
1	58
2	25
3	11
4	4
5 (or more)	2

Thus even when single men and families without dependent children were excluded from consideration, only 11 per cent of the remainder had the commonly accepted number of children while 83 per cent had fewer and 6 per cent more.

An analysis by a Belgian engineer, M. Midol, of the family status of over 11,000 employers in some coal mines, stone quarries, cement factories and construction shops of Belgium throws interesting light upon the family composition in that country and his figures are given below:

TABLE IV

FAMILY STATUS OF 11,313 BELGIAN WORKMEN

Status as regards dependents	Total number of workmen	Per cent
Total	11,313	100
Single or not heads of household	4,986	44.1
Total married	6,327	55.9
Married but no children (under 14 years)	1,942	17.2
Married and 1 child	2,021	17.9
Married and 2 children	1,183	10.4
Married and 3 children	567	5.1
Married and 4 or more children	535*	4.7

*The totals for the married workers do not agree with the sum of the individual items. This is an error, however, in M. Midol's figures which it is impossible to ascertain.

[1] Dr. Perret, Repartition des Allocations Suivant Les Charges Familiales. *Compte-Rendu III Congrès National des Allocations Familiales* (1923) pp. 80–81.

It may perhaps be objected that the exclusion of families with more than five children and those where boarders and lodgers were taken, resulted in a distortion of the real family composition of the 25,000 families as a whole. This objection seems at first to be corroborated by the fact that the average size of the 11,156 families was 3.96 persons as compared with 4.88 persons for all the families as a whole. Upon analysis, however, it is found that the average number of children under fourteen for the normal families was 1.90, and of children under sixteen in all the families only 1.98,[1] or when allowance is made for the difference in age groups, approximately the same average number of dependent children. The difference in the gross size of the two sets of families was therefore caused by the presence in the larger group of children over 16 and of adults, and not by more children of the dependent ages. The figures for the 11,156 families may consequently be regarded as affording an approximately accurate picture of the degree to which the families as a whole had children of less than 14 years dependent upon them for support.

The second important investigation was that made by the Illinois Health Insurance Commission in 1918 under the direction of my colleague Professor H. A. Millis of the University of Chicago. The families in thirty carefully selected blocks in working class neighborhoods in Chicago were studied, and the relative number of members in each secured. Single men living as boarders or lodgers were not included. It should be noted that the study deals with the "members" of a family which is, of course, a much broader term than "dependent children" since it not only includes the younger children but also others who are living at home and who are gainfully employed, and in addition adult family dependents as well.

Table VII gives this material.[2]

Even upon the basis of "members" instead of children, 65.2 per cent of the families were found to consist of fewer than five persons while only 13.2 per cent were precisely of that size. Nearly 22 per cent of the families in turn included more than five members.

The final and largest investigation was that conducted by Miss Marie Obenauer for the United States Coal Commission and which was based upon the 1920 census schedules of coal miners in the anthracite fields. Miss Obenauer secured data in all for 128,853 miners, of whom more than half were foreign born: 83,877, or 65.1 per cent of the total, maintained homes, leaving 34.9 per cent that did not do so. Of

[1] Computed from data on the following pages of the *Eighteenth Annual Report of the Commission of Labor*, 242–43, 256–57, 568.

[2] Report of the Health Insurance Commission of the State of Illinois, p. 191.

TABLE VII

COMPOSITION OF 3,048 CHICAGO WORKING CLASS FAMILIES, 1918

Number of members of family	Number of families	Per cent of families
1	106	3.3
2	559	18.3
3	718	23.6
4	609	20.0
5	400	13.2
6	293	9.6
7	178	5.8
8	97	3.2
9	55	1.8
10	15	.5
11	12	.4
12 (or more)	6	.2
Total	3,048	100.0

those who were heads of households, the following percentages had the given number of children under 16 years living at home with them.[1]

Here 44.3 per cent of all the families had fewer than three children while only 15.9 per cent had the exact quota. 39.2 per cent on the other hand had more than three children.

Number of children under 16 years	Percentage of mine workers maintaining homes with specified number of children
0	13.1
1	14.4
2	16.8
3	15.9
4	13.4
5	10.6
6	7.4
7	4.3
8 (or more)	4.1

Miss Obenauer made an especially interesting comparison of the composition of the families of native born and foreign born miners. 14.6 per cent of the native born had three children under the age of sixteen as compared with 16.6 per cent of the foreign born. The interesting differences occurred in the proportion which had fewer and more than three children respectively. 58.2 per cent of all the native born families had less than three children while only 36.2 per cent of the foreign born had less than this number. 39.2 per cent of the native born had indeed less than two children as contrasted with 20.8 per cent of the foreign born. The comparison as regards families with more than three children is perhaps even more striking. No less than 47.6 per cent of the foreign born exceeded this number as compared with 27.1

[1] Report of United States Coal Commission as given in *Weekly News Letter* of the Illinois State Federation of Labor, July 21, 1923, p. 3. I am indebted to Miss Obenauer and Dr. Helen M. Wright for clarifying many of the details of the study.

and single, out of every thousand, classified according to the number of children under 14 years dependent upon them for support was as follows.[1]

Number of dependent children under 14 years	Number of workmen out of 1,000 with given number of dependents
0	531
1	162
2	125
3	84
4	55
5	28
6	10
7	4
8 or more	1

This showed that only 8.4 per cent had three such children dependent upon them, while 81.8 per cent had less than this number and 9.8 per cent had more. Over 50 per cent had no children at all.

As a result of the general interest as to the relative number of dependents in the families, the British census in 1921 added a question calling for the number of children under 16 years of each married man or woman. When this was added to the statistics gathered concerning the number of single men and women in comparison with the number married, widowed or divorced, there was made possible for the first time a picture of the family composition of the households of millions of workers. The British census returns are being published separately by counties and the combined results for the country as a whole have not yet been issued. The returns for the first fourteen counties covering approximately 6.8 million gainfully occupied males of 20 years and over which are given below are, however, without doubt representative of conditions in general.

The following table shows the number of single, married, widowed and divorced adult males who were gainfully employed and the proportion which each formed of the total.[2]

Slightly over 26 per cent of all the adult males who are gainfully occupied are therefore unmarried. Only an insignificant proportion of these have children to support although more doubtless, have adults who are dependent upon them. Few, however, have the equivalent of a family of five to care for and it seems probable that the majority have no one completely dependent upon them. A wage sufficient to

[1] A. L. Bowley, "Earners and Dependents in English Towns," *Economica*, May, 1921 (Vol. 1), p. 111.

[2] The number of occupied males who were either married, or widowed and divorced is given directly for each county. The total number of males of 12 years and over is also given. By subtracting the number of males of the years 12–19 inclusive less the number of married juveniles from this figure, one secures the number of adult unmarried males of 20 years and over who are gainfully employed.

TABLE II

MARITAL CONDITION OF 6.8 MILLION ADULT OCCUPIED MALES (20 YEARS AND OVER) IN 14 ENGLISH AND WELSH COUNTIES—1921

County	Number of occupied males 20 years and over according to marital condition (in 000's)				Percentage of occupied males 20 years and over according to marital condition		
	Single	Married	Widowed or divorced	Total	Single	Married	Widowed or divorced
1. London	324	821	55	1,200	27	68	5
2. Middlesex	69	240	12	321	21	75	4
3. Durham	127	272	19	418	30	65	5
4. Essex	100	277	16	393	25	71	4
5. Glamorgan	107	234	17	358	30	65	5
6. Hertford	22	62	3	87	26	70	4
7. Kent	80	211	13	304	26	70	4
8. Lancaster	367	928	67	1,362	27	68	5
9. Monmouth	38	85	6	139	29	66	5
10. Stafford	93	256	17	366	25	70	5
11. Surrey	61	172	9	242	25	70	5
12. Warwick	94	272	17	383	24	71	5
13. Worcester	28	76	5	109	25	70	5
14. Yorkshire	305	817	55	1,177	26	69	5
	1,815	4,723	311	6,849	26.5	69.0	4.5

maintain a family of five would therefore be vastly in excess of the present needs of this group, comprising as it does over one-quarter of all the adults who are at work.

The married and widowed workers in each county are also classified according to the number of children under 16 that they have, as is shown in Table III. Due to the newness and complexity of the question calling for the number of children, it is probable that a consider-

TABLE III

RELATIVE NUMBER OF CHILDREN UNDER 16 YEARS PER FAMILY IN ENGLAND AND WALES—1921

County	Percentage of married men or widowers with varying number of living children under sixteen										
	No children or no statement	1	2	3	4	5	6	7	8	9	10 or more
1. London	43.8	22.6	15.0	8.7	5.0	2.8	1.4	.5	.2	0	0
2. Middlesex	42.7	24.0	16.2	8.6	4.4	2.3	1.1	.5	.2	0	0
3. Durham	34.6	22.9	17.0	11.4	7.1	4.1	2.0	.7	.2	0	0
4. Essex	42.2	23.0	15.7	9.0	5.1	2.8	1.4	.5	.2	.1	0
5. Glamorgan	35.8	23.1	16.6	10.7	6.6	3.9	2.0	.9	.3	.1	0
6. Hertford	45.7	22.7	15.0	8.0	4.5	2.3	1.1	.5	.2	0	0
7. Kent	46.0	22.5	14.7	8.1	4.4	2.4	1.2	.5	.2	0	0
8. Lancaster	42.5	23.9	15.5	8.8	4.9	2.6	1.2	.5	.1	0	0
9. Monmouth	34.4	23.3	17.0	11.1	6.8	4.0	2.1	.9	.3	.1	0
10. Stafford	38.0	23.1	16.2	10.1	6.2	3.6	1.8	.7	.2	.1	0
11. Surrey	46.5	23.0	15.1	7.8	4.0	2.1	1.0	.4	.1	0	0
12. Warwick	40.9	23.9	16.1	9.1	5.1	2.8	1.3	.5	.2	.1	0
13. Worcester	43.3	23.0	15.1	8.9	5.0	2.7	1.3	.5	.2	0	0
14. Yorkshire	43.1	23.7	15.3	8.7	4.8	2.6	1.2	.5	.1	0	0
Total	41.9	23.3	15.6	9.1	5.2	2.9	1.4	.5	.2	0	0

able proportion of parents did not make full reports and that an appreciable number of those who actually had children did not make any such statement. The British census, however, believes that a "reasonable degree of confidence" may be placed in the results for the married people, "accounting as they do for about 90 per cent of the children."[1]

The fact that they roughly check with Bowley's earlier study of ten thousand workers gives them a further degree of corroboration.

Forty-two per cent of the married workers would seem, therefore, to have had no dependent children at the time that the inquiry was made;[2] 23 per cent had only one child and 16 per cent had two, 81 per cent in all, therefore, had less than the number commonly set as the standard. Only 9 per cent of those that were married had the standard of three children while 10 per cent had more than this number.

Since the married and widowed workers formed only 73.5 per cent of all the adult males, it will be seen that the so-called standard family was in reality characteristic of only 7.6 per cent of all the workers. The average number per family of children under sixteen years in the fourteen counties was indeed only 1.31.

2. FRANCE AND BELGIUM

There have been two interesting investigations into the size of French families. The first was that conducted in 1919 by a group of Parisian manufacturers into the family status of 99,600 workmen who were almost exclusively employed in the metallurgical and metal industries of the capital. They found that 36.9 per cent of the total were either unmarried or widowed or divorced without having dependent children. On the other hand 63.1 per cent were married.[3]

These married workmen in turn had the following number of children dependent upon them:

Number of children	Per cent of married workmen
0	38.4
1	33.6
2	17.5
3	6.4
4	2.5
5 (or more)	1.6
Total	100.0

[1] See report on *County of Middlesex*, Census of England and Wales, 1921 (1923) p. XXXI. Similar statements are made in the case of most of the other counties.

[2] This is probably something of an overestimate since it includes those who made no statement but yet had dependent children.

[3] See Pierre Richemond, Allocations Pour Charges de Famille et Caisses de Compensation, *Revue D'Economie Politique*, Vol. XXIV, pp. 590–91.

The family of three children thus seemed to be characteristic of only 6.4 per cent of the 63 per cent of married workers or only 4 per cent of the total number of workers, while less than this number had more than three children dependent upon them for support.

The second investigation was that made in 1923 by Dr. Perret of Lyons for the Comité des Allocations Familiales and which covered several hundred thousand employees. Dr. Perret found that only 24.7 per cent of the workers were fathers of children.[1] Since a large number of women and juveniles, however, were included as workers, this figure should not be greatly relied upon. The average number of children in families where there were any at all was 1.66 and the proportion which had the varying number of children was as follows:

Number of children in family	Per cent of families
1	58
2	25
3	11
4	4
5 (or more)	2

Thus even when single men and families without dependent children were excluded from consideration, only 11 per cent of the remainder had the commonly accepted number of children while 83 per cent had fewer and 6 per cent more.

An analysis by a Belgian engineer, M. Midol, of the family status of over 11,000 employers in some coal mines, stone quarries, cement factories and construction shops of Belgium throws interesting light upon the family composition in that country and his figures are given below:

TABLE IV

FAMILY STATUS OF 11,313 BELGIAN WORKMEN

Status as regards dependents	Total number of workmen	Per cent
Total	11,313	100
Single or not heads of household	4,986	44.1
Total married	6,327	55.9
Married but no children (under 14 years)	1,942	17.2
Married and 1 child	2,021	17.9
Married and 2 children	1,183	10.4
Married and 3 children	567	5.1
Married and 4 or more children	535*	4.7

*The totals for the married workers do not agree with the sum of the individual items. This is an error, however, in M. Midol's figures which it is impossible to ascertain.

[1] Dr. Perret, Repartition des Allocations Suivant Les Charges Familiales. *Compte-Rendu III Congrès National des Allocations Familiales* (1923) pp. 80–81.

Thus here again the so-called standard family is found in practice to be characteristic of approximately only 5 per cent of all the workers while a full 90 per cent have fewer than three children to support.

3. INDIA

It might naturally be expected that in the prolific Orient, the family of three children would be the rule and not the exception. An investigation by G. Findley Shirras into 3,073 working-class budgets in Bombay shows, however, that even here such a family is distinctly in the minority.[1] 603 of these budgets, or 19.6 per cent of the total were those of single men leaving 2,473, or 80.4 per cent of the total, those of families. The families fell into the following groups as regards the number of children under fourteen in the family.[2]

TABLE V

COMPOSITION OF 2473 WORKING CLASS FAMILIES IN BOMBAY 1921-22

Number of children under fourteen in family	Number of families with the given number of children	Percentage of families with the given number of children
1	253	10.2
1	659	26.7
2	845	34.2
3	462	18.7
4 (or more)	254	10.2
Total.....	2,473	100.0

Thus 71.1 per cent of the Indian workmen with families had less than three children while only 18.7 per cent had this "standard" number. Slightly over 10 per cent in turn had more than this number.

These results need to be qualified, however, in two ways. First, an allowance should be made for adult dependents living with the family in Bombay. In approximately 13 per cent of the cases there were two or more adult women in the family circle and some of these were doubtless dependents. Second, a large number of the dependents of these workers had not been brought to the city but were left behind in the country. These, unfortunately, were not classified according to the size of the family but they amounted on an average to three-fifths of a person for every worker and were evenly divided as between men, women and children under fourteen.[3]

[1] G. Findley Shirras, *Report on an Enquiry into Working Class Budgets in Bombay*, Labour Office, Government of Bombay, 1923.

[2] *Ibid.*, p. 9.

[3] *Ibid.*, p. 7.

4. THE UNITED STATES

The 1920 census of occupations showed that 27.6 per cent of all the males between the ages of 20 and 65 were unmarried and this proportion may be accepted as applying to all gainfully employed adult males.[1] Since the number of gainfully employed adult males in the same age group was approximately 28.3 million in 1920, we may then estimate the number of unmarried male workers of from 20 to 65 years of age at 7.8 millions. These certainly do not have families of five to support.

While the census schedules contain data on the size of families, this material has never been tabulated or made available in the official reports. There are three investigations, however, which do give admirable material upon this very question. The first of these is contained in the well-known *Eighteenth Annual Report of the United States Bureau of Labor*.[2] This investigation secured the budgets of 25,440 working class families in 33 states and as a by-product of its work collected material on the family composition of 11,156 of these which were classified as "normal." A "normal" family was defined as one where the husband was at work, where there was a wife and not more than five children and none over fourteen and finally where there was no "dependent,[3] boarder, lodger or servant."

The following table shows the number and proportion of families with varying numbers of children.[4] All bachelors, it should be remembered, are excluded from consideration and only families therefore are covered.

TABLE VI

FAMILY COMPOSITION OF 11,156 "NORMAL" FAMILIES IN 1901

Number of children under 14	Number of families	Percentage of families
0	2,124	19.0
1	2,579	23.1
2	2,700	24.2
3	1,973	17.7
4	1,248	11.2
5	532	4.8
Total	11,156	100.0

The family of three children was therefore characteristic of only 17.7 per cent of the total number. Sixteen per cent of the families had more than three children while 66.3 per cent, or virtually two-thirds had less.

[1] Computed from dates given in *Abstract of the Fourteenth Census*, p. 222.

[2] "The Cost of Living and Retail Prices of Food."

[3] *Ibid.*, p. 90. By this was undoubtedly meant an adult dependent distinct from the wife.

[4] *Eighteenth Annual Report*, U. S. Commission of Labor, p. 568.

It may perhaps be objected that the exclusion of families with more than five children and those where boarders and lodgers were taken, resulted in a distortion of the real family composition of the 25,000 families as a whole. This objection seems at first to be corroborated by the fact that the average size of the 11,156 families was 3.96 persons as compared with 4.88 persons for all the families as a whole. Upon analysis, however, it is found that the average number of children under fourteen for the normal families was 1.90, and of children under sixteen in all the families only 1.98,[1] or when allowance is made for the difference in age groups, approximately the same average number of dependent children. The difference in the gross size of the two sets of families was therefore caused by the presence in the larger group of children over 16 and of adults, and not by more children of the dependent ages. The figures for the 11,156 families may consequently be regarded as affording an approximately accurate picture of the degree to which the families as a whole had children of less than 14 years dependent upon them for support.

The second important investigation was that made by the Illinois Health Insurance Commission in 1918 under the direction of my colleague Professor H. A. Millis of the University of Chicago. The families in thirty carefully selected blocks in working class neighborhoods in Chicago were studied, and the relative number of members in each secured. Single men living as boarders or lodgers were not included. It should be noted that the study deals with the "members" of a family which is, of course, a much broader term than "dependent children" since it not only includes the younger children but also others who are living at home and who are gainfully employed, and in addition adult family dependents as well.

Table VII gives this material.[2]

Even upon the basis of "members" instead of children, 65.2 per cent of the families were found to consist of fewer than five persons while only 13.2 per cent were precisely of that size. Nearly 22 per cent of the families in turn included more than five members.

The final and largest investigation was that conducted by Miss Marie Obenauer for the United States Coal Commission and which was based upon the 1920 census schedules of coal miners in the anthracite fields. Miss Obenauer secured data in all for 128,853 miners, of whom more than half were foreign born: 83,877, or 65.1 per cent of the total, maintained homes, leaving 34.9 per cent that did not do so. Of

[1] Computed from data on the following pages of the *Eighteenth Annual Report of the Commission of Labor*, 242–43, 256–57, 568.

[2] Report of the Health Insurance Commission of the State of Illinois, p. 191.

TABLE VII

COMPOSITION OF 3,048 CHICAGO WORKING CLASS FAMILIES, 1918

Number of members of family	Number of families	Per cent of families
1	106	3.3
2	559	18.3
3	718	23.6
4	609	20.0
5	400	13.2
6	293	9.6
7	178	5.8
8	97	3.2
9	55	1.8
10	15	.5
11	12	.4
12 (or more)	6	.2
Total	3,048	100.0

those who were heads of households, the following percentages had the given number of children under 16 years living at home with them.[1]

Here 44.3 per cent of all the families had fewer than three children while only 15.9 per cent had the exact quota. 39.2 per cent on the other hand had more than three children.

Number of children under 16 years	Percentage of mine workers maintaining homes with specified number of children
0	13.1
1	14.4
2	16.8
3	15.9
4	13.4
5	10.6
6	7.4
7	4.3
8 (or more)	4.1

Miss Obenauer made an especially interesting comparison of the composition of the families of native born and foreign born miners. 14.6 per cent of the native born had three children under the age of sixteen as compared with 16.6 per cent of the foreign born. The interesting differences occurred in the proportion which had fewer and more than three children respectively. 58.2 per cent of all the native born families had less than three children while only 36.2 per cent of the foreign born had less than this number. 39.2 per cent of the native born had indeed less than two children as contrasted with 20.8 per cent of the foreign born. The comparison as regards families with more than three children is perhaps even more striking. No less than 47.6 per cent of the foreign born exceeded this number as compared with 27.1

[1] Report of United States Coal Commission as given in *Weekly News Letter* of the Illinois State Federation of Labor, July 21, 1923, p. 3. I am indebted to Miss Obenauer and Dr. Helen M. Wright for clarifying many of the details of the study.

per cent of the native born. With the great restriction of immigration and the Americanization of the children of these miners, the birth rate may confidently be predicted to approach more closely that which at present prevails for the native born, with a consequent decrease in the relative importance of families with three or more children simultaneously under the age of sixteen and an appreciable increase in the families with less than three children.[1]

As a matter of fact, therefore, taking the country as a whole, the adult male workers do not on the average have four additional persons dependent upon them. Since there were approximately 28.3 millions such males gainfully employed in 1920, to have given each enough to have maintained a family of five would have meant providing for slightly over 141 millions of people, whereas there were only slightly less than 106 millions in the United States as a whole at that time. Payment would therefore have been made for 35 million fictitious wives and children even were the adult males to be considered the sole supporters of the people of the country. But there were no less than 6.2 millions of adult women and 6.3 millions of juveniles of from 16 to 20 years who were also employed. Even granting that these just earned enough to maintain themselves, we would then have only between 93 and 94 millions of people to be supported by the 28.3 million adult males who were gainfully employed. To pay each of these adult males enough to maintain a family of five, while continuing to pay women and children as before would amount to paying for approximately forty-eight million people that did not exist. To adopt the principle of the feminists and pay the gainfully employed women the same amounts as the men would mean providing maintenance for no less than seventy-two million fictitious persons.

There were indeed in 1920, 33.6 million children under fifteen years of age and 32.2 million males of 20 years and over.[2] This amounted to only 1.04 children for every adult male. Even if the full burden of support of these children fell upon the 28.3 million occupied males between the ages of 20 and 65 there would be only 1.19 children for every such male worker.

There were in 1920 4.9 millions of persons who were 65 years of age or over, or only .17 such persons for every gainfully employed adult

[1] Another investigation that is worthy of mention is that conducted by Prof. Jessica B. Peixotto and Prof. Solomon Blum of the University of California into the composition of the families of eighty-two members of the typographical union of San Francisco. Single men were excluded. Twenty-six of the eighty-two were found to have no children, twenty-two to have one child, twenty to have two children, six to have three, six to have four and only one to have five and six children respectively. Over 82 per cent, therefore, had less than three children, and only 7 per cent had precisely that number. Between 11 and 12 per cent had more than three children. I am indebted to Miss Peixotto and to Mr. Blum for permission to use these figures.

[2] *Fourteenth Census*, Vol. III, p. 16.

male who was less than 65 years old.[1] There were moreover .72 wives which with the other two groups would amount to 2.08 persons who might be regarded as dependent on the average upon the male adult worker. To this should be added the burden of the support of children of fifteen and over and of dependent adults of less than 65 years. At the most however the total could not amount to more than 2.35 dependent persons.

5. WHY HAS THE FAMILY OF FIVE BEEN ACCEPTED AS THE STANDARD?

The preceding analysis has seemed to destroy the validity of the "standard" family of five as the most representative type. In virtually every investigation that has been made such a family has been disclosed as characteristic at any one time of only a relatively small percentage of the total cases.

How then may these results be reconciled with the fact that a large number of investigations such as that of Chapin in 1908–9 in New York City and the country-wide study of the Bureau of Labor Statistics in 1917–18, have shown that working class families do in fact average not far from five members? There are several reasons for this apparent divergence. The first is that virtually all these studies have excluded single men from consideration and that these form between one-quarter and three-tenths of all the adult male workers. The second is that these studies have frequently counted all members of the family whether actually dependent or not and have included sons and daughters who were really self-supporting. These figures on the size of families have in turn been illogically taken over as a basis for computing dependency. A third reason is that in some of the investigations, notably those of Chapin and the 1918 study of United States Bureau of Labor Statistics, the families have been expressly selected in order that the results might conform approximately to the standard type.[2]

A final potent source of confusion has been the acceptance of the arithmetic average itself as the measurement of a family's composition or need. Families might average five members but only a small percentage of them might actually have such a number with the vast majority having either more or less. The delusive character of the arithmetic average has long been perceived for most branches of social

[1] *Abstract of the Fourteenth Census*, p. 138.

[2] This was not done to deceive the public. These investigations merely accepted the family of five as the standard and centered their efforts on finding out how such families spent their money and how much they needed to live. Some, however, have forgotten the method of sampling employed and have cited these studies as proof that five constitutes the representative American workingman's family.

phenomena but its adequacy in matters of family composition has not been challenged to the extent that it should have been.[1]

6. LARGE NUMBERS OF CHILDREN ARE CONTAINED IN THE FAMILIES WITH MORE THAN FIVE MEMBERS

There is one other point that needs to be noted and that is the fact that while only a small percentage of the families, and consequently a still smaller proportion of the adult workmen have more than three dependent children at any one time, a much larger percentage of the total number of children are included in these families. Thus in the fourteen English and Welsh counties only 10 per cent of the married or widowed workmen (who in turn formed only 73 per cent of the adult male workers) had more than three children. The children in these families, however, comprised between 37 and 38 per cent of all the children. It will be remembered that the 1901 investigation by the United States Bureau of Labor showed that families with four and five children formed only 16 per cent of the total number. Thirty-seven per cent of all the children were, however, found in these very families.

To grant a wage to all sufficient to maintain a family of five would, therefore, entail much more suffering than would be apparent at first sight. The important feature to be considered in such a situation would not so much be that 10 to 15 per cent of the families would have a deficiency in their budgets as that from 35 to 40 per cent of the children would lack some of the essential things of life.[2]

7. SUMMARY AND A FINAL SUGGESTION

The preceding analysis may be summarized as follows:

1. The burden of dependents is not uniform upon all wage-earners but varies widely.

2. Only a small percentage of adult male workers actually have the standard family of five to maintain. In England not more than 10 per cent of the workers and possibly even less have such families. In the United States such evidence as is available indicates that probably not more than 12 per cent of the workers have such a family.[3]

[1] Miss Margaret L. Stecker's excellent article, "Family Budgets and Wages," *American Economic Review*, Vol. 11, pp. 447–65, is not as strong an assault upon the family of five concept as the norm for measuring wages as it would have been had she emphasized the inadequacy of any average as furnishing a standard for all.

[2] This important consideration is emphasized by B. Seebohm Rowntree in *The Human Needs of Labor*, pp. 38–39. From an analysis of census schedules for the city of York, Mr. Rowntree shows that a wage that would be insufficient for 33 per cent of the families for one year or more would be insufficient for 62 per cent of the children for an equal period of time.

[3] That is when allowance is made for bachelors.

3. The majority of the workers have fewer dependents than a wife and three children. In England approximately 26 per cent of the adult male workers are bachelors, while probably at least two-thirds of the married have either no children at all or only one or two. In the United States approximately 28 per cent of the gainfully employed adult males are unmarried. Somewhere between 50 and 60 per cent of all the married workers had less than three children.[1]

4. A considerable percentage of the workers have more than three children. In England, probably somewhere around 10 per cent of the workers have four or more children to support. In the United States the proportion is apparently higher and may reach 15 per cent. The proportion of the total number of children that are in these families is of course much higher. The evidence seems to indicate that 37 per cent of the English children come from this group of families and such evidence as we have that approximately 40 per cent or more of the children of our industrial wage-earners come from such families. It is these families upon which the burden of need is heaviest and the chief sufferers are the children.

5. The statistics for other than English speaking countries indicate that the same general tendencies that have been stated apply there although the details naturally vary.

6. To pay all workers a wage sufficient to maintain a family of five would be just adequate for only a small percentage of the workers. It would be more than adequate in English speaking countries for between 70 and 80 per cent of all workers while it would be less than was needed in from 10 to 15 per cent. The granting of such a wage in the United States would mean paying for no less than 47 million fictitious dependents. Combined with equal pay for women, it would mean paying for no less than 72 million persons who did not exist. Without meeting the basic needs of those with large families, it would pour into the pockets of the unmarried and the married with few or no children a great excess over their wants.

In the light of these facts, it seems difficult to justify the family of five as the standard by which the wages of all adult males should be fixed. Advocates of the living wage doctrine, a principle which the author believes to be ethically and economically sound, do their cause a disservice by tying it up with so fallacious a measuring stick. Should not the real principle be that as needs are not uniform but variable so minimum wages should not be uniform but should vary according to the needs of the worker and his family? In other words should not the

[1] The low percentage among the foreign-born anthracite coal miners is probably not characteristic of the American working class population as a whole.

single man receive enough to maintain him plus a comfortable margin to permit for saving for marriage, but not enough to support a non-existent family of five? Then as his responsibilities increased, so would his allowance. Under such a system, the added cost occasioned by a dependent wife, or by dependent children would be met by allowances added to the basic wage of the worker.

Those with large families would be protected in a manner impossible under any system of a uniform minimum wage while those with few or no dependents would not be given an unneeded surplus. The many complex problems raised by this proposal, the present writer hopes to treat at length in a forthcoming book.[1]

[1] I have already discussed some aspects of these problems in "Factors in Wage Determinations," Papers and Proceedings of the Thirty-Fifth Annual Meeting of the American Economic Association Supplement to the *American Economic Review*, Vol. XIII, p. 141; "Wages, Regulation and Childrens' Maintenance in Australia," *Quarterly Journal of Economics*, August, 1923, Vol. XXXVII, pp. 644–86; "Family Allowances and Clearing Funds in France," *Quarterly Journal of Economics*, February, 1924. Vol. XXXVIII.

A NEW CLEARINGS INDEX OF BUSINESS FOR FIFTY YEARS

By Carl Snyder, *Federal Reserve Bank of New York*

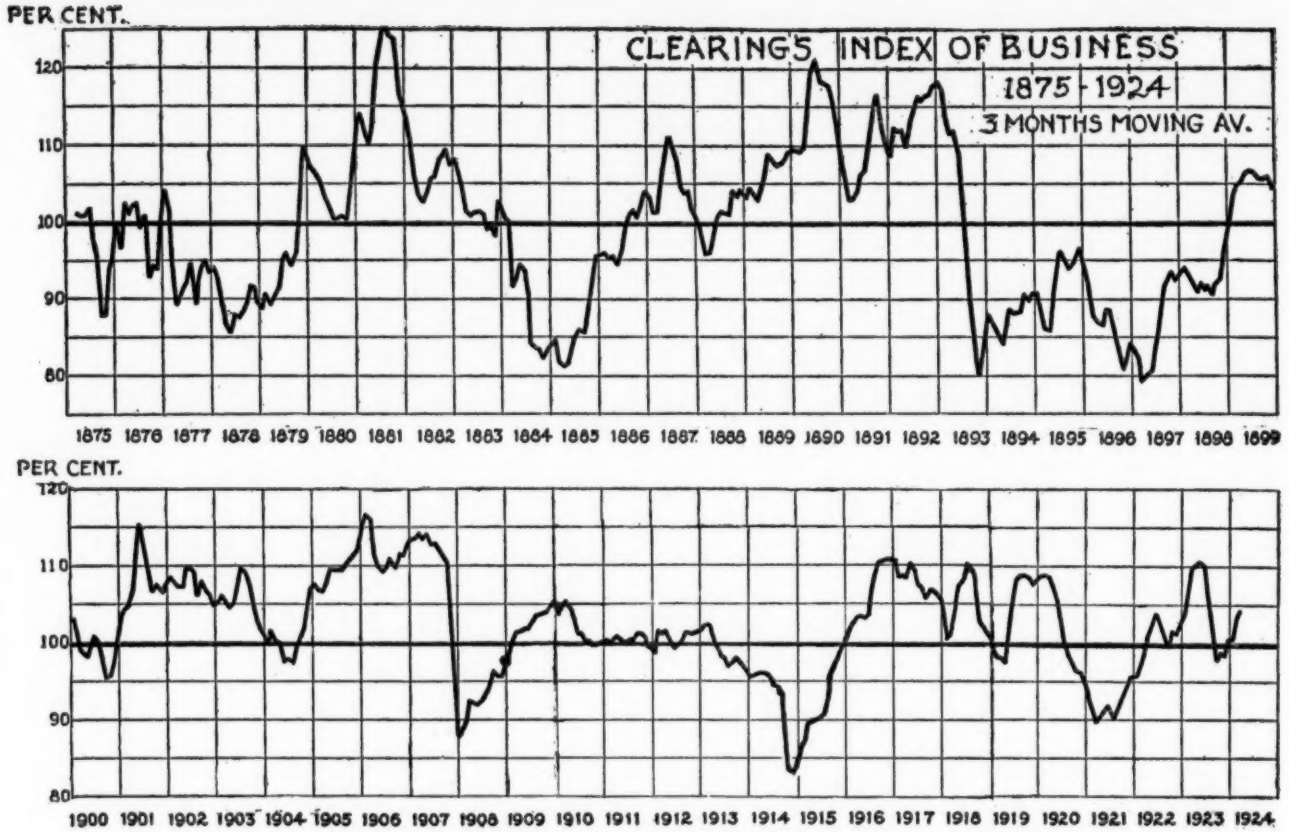
In the June issue of this Journal I described a new index of the General Price Level, as contrasted with indexes of commodity prices at wholesale, or of retail prices, wages, rents, and the like. This General Price Level was defined as the average of all payments entering into bank clearings, since in this country some 80 per cent of all exchanges are made through bank checks.

The tests proposed for the weighting of the different components entering into this new index were two: first, that the abnormal figures for bank clearings, or debits, after 1915, due to the extraordinary rise of prices in the war, and their subsequent collapse, should by means of this index be brought down to such a level as would continue through the war period and after the same trend of growth as that of the preceding thirty years; second, that the deviations from this trend of clearings thus corrected should closely correspond in time and in percentages with our composite index of fifty-six independently computed series which we have called the Volume of Trade.

Both of these tests, it was noted, were easily fulfilled since an empirical weighting of the components of the new index of the General Price Level had produced a result corresponding closely to these two requirements, so that it needed only a little shifting of the weights to produce a very close fit with the index of the Volume of Trade. Secure in this result it was then possible, as was therein noted, to make use of bank debits, or bank clearings, as a measure of business for a prolonged period backward, with a certainty which had never before been attainable.

In other words, if it was possible, in a period of such tumultuous change in price levels as that from 1916 to 1924, so to correct bank clearings as to make these a measure of business changes as accurate as a composite index from the broadest sampling of data now attainable, it was still easier to do this in the antecedent period when changes in the general price level were slow and relatively small. The present note is to describe this new index of the Variations in Trade through the last fifty years.

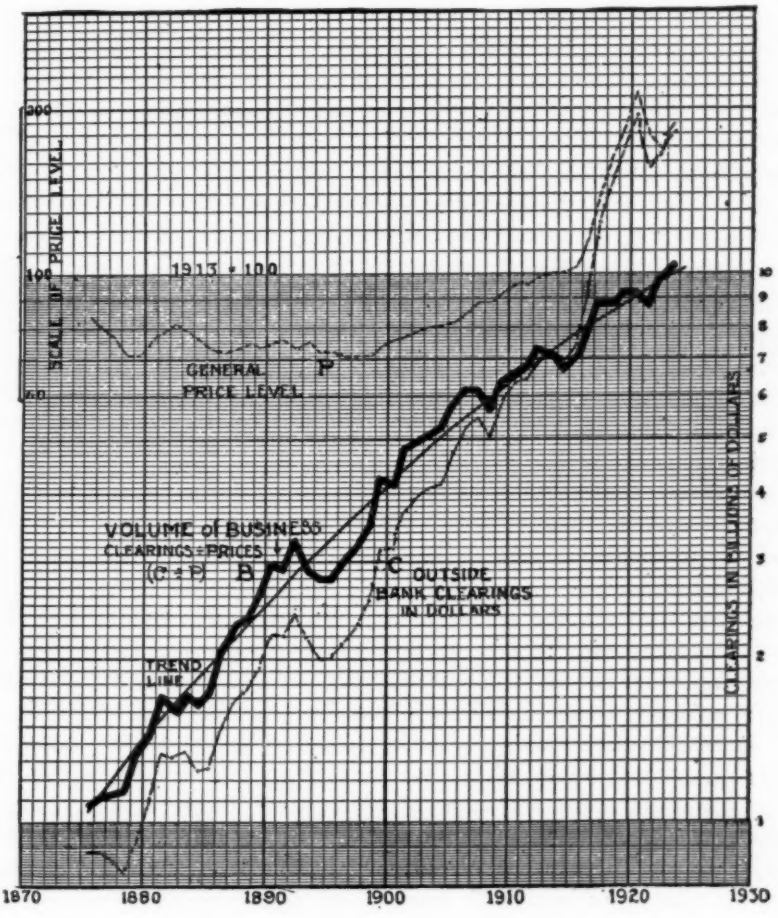
First, as to the method of construction. In the accompanying chart there is shown, first of all, bank clearings outside of New York City,



BUSINESS ACTIVITY OVER 50 YEARS, AS REFLECTED IN BANK CLEARINGS OUTSIDE NEW YORK CITY. CORRECTED FOR PRICE AND SEASONAL VARIATIONS

(Computed Trend = 100 Per Cent)

in dollars, (line C), exactly as reported from month to month throughout the period. With this is given above the new index of the General Price Level (line P). It will be observed that when the figures for outside bank clearings are divided by the index numbers of the General Price Level (C÷P), the irregular line C is transformed into a curve of very definite trend and characteristic (line B), through which has been put a parabolic curve of growth, as shown. Taking the points month by month on this line of growth as a base of 100, the deviations of the



BANK CLEARINGS CORRECTED FOR PRICE CHANGES

The heavy line, representing volume of business (B), is obtained by dividing actual bank clearings outside of New York (C) by a new index of the General Price Level (P). See text. The trend line shown is a computed curve.

corrected line B are then calculated in percentages of these points of normal growth; and these percentages constitute the new index.

Why do we feel justified in believing that this is an accurate measure of business variations throughout the period?

Plainly this corrected line of clearings (line B) is not an accurate index of the growth of trade in all this period, for the whole country. In the earlier half, from 1875 to about 1900, it rises at a much higher rate than, for example, the composite index of production in leading commodities, presented at Atlantic City four years ago.[1] There the evidence seemed to indicate a fairly even and continuous rate of growth throughout the whole period, averaging not very far from 3½ per cent per annum. This result was confirmed by the parallel investigations of King, Stewart and Day for portions of this period. From about 1900 the rate of growth in trade, as indicated by bank clearings, is very near to that calculated as the average for industrial production.

The reason for this changing rate of increase is twofold: first, that the quarter of a century from 1875 to 1900 was the period in which our bank clearings system began, and attained practically its maximum position as a business register. In the first ten years of this period the number of clearing houses alone was doubled. And in the same quarter of a century undoubtedly there was a distinct growth in the use of checks as a means of payment. Naturally, then, the total of clearings probably rose much more rapidly than the actual volume of trade or exchanges.

But from about 1900 on there seems evidence that the use of checks had become, as we say, nearly universal, at least among corporations and business firms. Since that time, no doubt, the increase in the number of individual accounts has been far greater than the actual increase in trade, but these are for the most part small accounts of private depositors, which, against the colossal totals from the main avenues of trade, do not greatly affect the result. Finally, from about 1900 we have a fairly continuous number of reporting centers.

But while the rate of growth of clearings in the earlier half of the period was much more rapid than in the later half, it was still throughout the whole period even and consistent growth, not sensibly affected by business conditions. The amounts of clearings reported, from year to year, however, were deeply affected by the ups and downs of trade, and registered these closely; and there is reason to think that the deviations from the trend of growth are quite as accurate a picture of variations in actual business and trade in the earlier half as in the later. This view is confirmed by a comparison of the results from this

[1] Snyder: *The American Economic Review*, March, 1921, p. 70.

source with similar results from all other available data. The most valuable and reliable of these data is probably the new index of the Turnover of Bank Deposits which we have now worked out for the same period, and which will be described in a later issue. In a general way the results from this new index compare closely with those from the new Clearings Index, alike in percentages and time. The conformity is not strict but it is none the less so general, both in time relations and percentage relations as, I believe, to leave very little doubt that an accurate measurement of business throughout this period has now been attained.

This view is further supported by a comparison of the new index with the variations in pig iron production from the line of normal growth in that industry. Naturally the ups and downs of pig iron are much more violent and erratic than the variations in the volume of trade; but if the two sets of variations be reduced to a standard deviation, the correlation both in time and percentage relations is fairly high. This correspondence is, on the whole, closer than might have been expected. We know definitely from a comparison of iron production with our composite index of the Volume of Trade from 1919 that pig iron (and, in general, production in the producer type of goods) is not now a reliable index of general business. Both pig iron and its congeners are as a rule distinct laggards behind the general business movement and are much more erratic in their course. But, allowing for this usual lag, and for the much greater amplitude of swing, pig iron production for the whole period comes much closer to the general trend of business, as shown by the clearings index, than has been true of the five and a half years for which the full data were available for our composite index of the Volume of Trade. Does this suggest that iron was formerly a much better "barometer" of business than it is now?

At any rate we have two accessory indexes, one from the field of industrial activity, another from that of activity of bank deposits, which in general yield much the same picture as that derived from the volume of bank clearings. Unsupported, no one of the three indexes could perhaps be accepted as a reliable measure of business, although the *a priori* evidence for the clearings index, from its close correspondence with the Volume of Trade index from 1919, is very strong. But taken together the general correspondence of all three could scarcely be a haphazard result.

And, it may be noted, this is practically all the evidence we have, or ever shall have. We have other indexes for the whole period, as of variations in interest rates, and bond yields, and business failures, and

in part for building construction and some other data. But none of these could be taken as a measure of business variations. They provide interesting correlations but are not in themselves bases of departure.

If this then be the definitive picture, two or three observations may be permitted. The first is that the violence or amplitude of business fluctuation seems to have been, throughout the whole period, steadily, as the electricians would say, damping down. In the last quarter of a century we have had no crisis comparable in prolonged severity with the major crises of '73 and '93. And on this picture even the collapse of 1921 appears to have been no more severe, measured in actual trade, than that of 1907.

There was in 1920–21 a violent fall in prices and the losses sustained by individual business firms were enormous. But this did not seem to involve corresponding decline in the actual volume of goods and services exchanged. In other words, the evidence seems to be that the ups and downs of trade in the aggregate which we have come vaguely to call the business cycle are steadily being ironed out with the development of the country, the integration of industry, and perhaps with the steady growth of larger business knowledge and therefore vision. These periods seem also steadily shortening in duration.

The next seemingly clear inference appears to be that there is no definite period or cycle. Indeed, an attempt to define or number the cycles which have been completed in this period would be largely arbitrary and a matter of choice.

Finally, it would appear that the fluctuations of trade have little or no relation to changes in the general level of prices or average of all payments. Interesting relationships of certain types and groups of commodities to the "cycle" have been worked out by Professor Persons, Professor Fisher and others; but with the general level of all prices there seems to be none. The widest fluctuations took place in a period of relatively small change in the general price level; while in the period of the most violent price change the oscillations of trade appear to have been among the smallest. This latter result is one of great significance for those who have interested themselves in the problem of a more stable system of purchasing media.

NOTE

The index numbers for the new index of the General Price Level were given in the June number of this JOURNAL. Those for the new Clearings Index of Business, in terms of a three-months moving average, are appended. Both of these indexes were largely the work of Miss Irene C. Sheehan, of the Reports Department, Federal Reserve Bank of New York.

CLEARINGS INDEX OF BUSINESS

Computed trend=100 per cent

Bank clearings, outside of New York City, 1875–1918; debits outside of New York City, 1919–1924; 3 months' moving average

Seasonal variation and price changes allowed for

YEARS	JAN.	FEB.	MARCH	APRIL	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.	AVER.
1875			101.1	100.8	100.9	101.6	98.0	95.3	87.9	88.0	94.0	95.1	96.3
1876	99.9	96.8	102.4	101.0	102.0	102.3	99.2	100.6	92.9	94.4	93.9	99.1	98.7
1877	104.0	101.1	95.2	89.4	91.7	92.2	94.6	89.4	92.6	94.1	95.1	93.5	94.4
1878	94.1	92.7	89.5	86.8	85.8	88.0	87.8	88.4	89.6	91.9	91.5	89.6	89.6
1879	88.9	90.4	89.5	90.4	91.8	95.2	96.0	94.3	95.6	100.9	106.7	109.9	95.8
1880	106.9	106.7	106.1	105.4	103.6	102.1	100.3	100.4	100.9	100.3	103.3	108.2	103.7
1881	112.9	114.2	111.2	110.2	112.3	120.3	123.4	126.1	124.2	123.9	119.3	116.3	117.9
1882	113.3	110.4	105.6	103.1	102.8	104.3	105.6	105.9	108.2	108.3	109.5	107.4	107.0
1883	108.1	106.3	104.8	101.3	100.9	101.3	101.47	101.0	99.2	99.7	98.1	102.9	102.1
1884	100.4	100.0	91.9	92.7	94.5	93.6	90.2	84.2	83.6	83.5	82.3	83.5	90.0
1885	84.5	84.8	81.9	81.1	81.7	84.9	85.7	86.1	85.7	87.5	91.9	95.6	86.0
1886	95.7	95.7	95.2	95.2	94.5	96.1	99.4	101.2	101.46	100.7	102.5	104.0	9847
1887	103.3	101.2	101.2	104.3	108.3	111.46	110.1	108.2	104.6	103.8	104.0	101.8	105.2
1888	100.0	98.2	95.7	95.9	97.3	100.4	101.4	101.1	101.0	104.2	103.4	104.4	100.3
1889	103.2	104.7	104.0	103.0	104.8	105.9	108.9	107.7	107.3	107.5	108.1	109.2	106.2
1890	109.4	109.3	109.1	109.7	115.7	119.4	121.1	118.1	118.0	117.7	116.0	112.7	114.7
1891	108.8	105.9	102.9	103.1	104.2	106.3	106.6	110.9	115.3	116.6	112.4	109.9	108.6
1892	108.7	112.5	111.9	112.1	109.9	113.4	114.7	116.5	115.9	116.6	116.7	117.7	113.9
1893	118.2	117.3	113.7	111.6	111.7	108.9	102.2	89.5	82.6	80.0	83.2	85.7	100.4
1894	88.0	86.4	85.1	84.1	87.4	88.7	88.1	88.1	88.2	90.4	89.9	90.9	87.9
1895	91.0	88.2	86.0	85.7	90.9	94.0	96.4	95.0	94.0	94.46	95.0	96.5	92.3
1896	93.7	91.8	87.8	87.0	86.7	88.8	88.5	85.3	82.3	80.6	81.9	84.2	86.6
1897	83.2	82.3	79.6	80.3	80.9	83.46	85.6	88.2	91.9	92.6	93.5	92.7	86.2
1898	93.3	94.2	93.1	91.8	91.0	92.2	91.2	91.6	90.4	91.9	92.7	96.1	92.5
1899	100.4	103.4	104.9	105.2	106.4	106.6	106.5	105.8	105.6	105.9	105.1	104.46	105.0
1900	103.1	101.3	99.0	98.3	99.46	100.7	100.1	98.2	95.3	95.6	97.1	100.3	99.0
1901	103.46	104.3	104.6	106.8	112.46	115.3	112.6	109.2	106.6	107.46	106.4	107.3	108.1
1902	108.5	108.1	107.3	107.4	109.7	109.7	109.2	106.1	108.0	106.8	106.2	104.9	107.7
1903	105.1	106.0	105.2	104.7	105.0	107.8	109.7	109.0	107.1	104.6	102.7	101.9	105.7
1904	100.0	101.5	100.1	100.0	97.6	98.0	97.5	99.4	100.8	101.9	104.6	107.0	100.7
1905	107.8	106.6	106.46	107.6	109.8	109.8	109.6	109.6	110.3	111.2	111.4	112.0	109.4
1906	115.8	116.46	116.0	111.4	109.9	109.2	109.6	110.9	109.7	111.5	111.1	112.4	112.0
1907	113.3	113.5	114.4	113.5	114.1	112.8	112.9	112.2	111.0	110.4	101.6	93.8	110.3
1908	87.6	89.9	92.6	92.4	91.9	92.3	93.6	94.2	96.3	95.6	95.8	96.4	93.2
1909	98.3	99.9	100.6	101.5	101.5	101.9	101.7	102.9	103.4	103.6	103.8	104.3	102.0
1910	104.4	103.9	104.8	105.6	104.9	103.1	101.46	101.7	100.8	100.2	99.9	100.0	102.6
1911	100.5	100.2	101.0	100.2	100.8	100.8	101.5	101.5	101.0	99.7	99.0	98.1	100.4
1912	98.6	101.4	101.3	101.8	100.8	100.0	99.4	99.8	100.7	101.9	101.3	101.3	100.7
1913	101.6	102.7	102.6	100.8	99.1	98.5	98.3	97.0	97.8	98.1	97.2	96.4	99.2
1914	95.7	96.4	96.2	96.1	95.1	94.9	94.7	93.3	90.1	85.4	83.1	83.0	92.0
1915	83.8	86.1	87.5	89.3	89.7	90.0	90.0	90.3	91.2	92.7	96.1	99.1	90.5
1916	100.2	102.2	102.9	103.4	103.3	103.4	103.5	104.4	106.1	108.8	110.3	110.4	104.9
1917	110.4	108.2	108.5	108.5	110.3	109.6	108.0	107.7	106.0	107.4	107.0	106.7	108.2
1918	104.1	100.4	101.1	103.7	107.5	108.0	108.4	110.3	110.0	109.0	103.5	101.4	105.6
1919	99.8	98.4	98.2	97.7	101.0	104.8	108.1	108.9	108.9	108.7	107.6	107.9	104.2
1920	108.6	108.8	108.7	107.3	105.6	103.0	100.3	98.4	97.4	96.3	96.1	95.2	102.1
1921	93.5	91.3	89.6	90.5	91.0	91.9	90.2	91.0	91.7	93.6	94.5	95.5	92.0
1922	95.7	97.0	98.6	101.1	102.8	104.0	102.7	101.1	99.6	101.6	101.3	102.2	100.6
1923	104.1	107.9	109.7	110.1	110.6	110.1	106.0	101.7	97.9	99.0	98.8	100.3	104.7
1924	100.3	103.47	104.4										

REVISION OF THE WEEKLY INDEX NUMBER

By Irving Fisher

In this Journal, September, 1923, I described "A Weekly Index Number of Wholesale Prices," published by me in the newspapers beginning with the first week of January, 1923.

It is my purpose to revise this index every year in the light of new data available. Probably the revisions will be slight, except every ten years when new Census figures make possible a general revision of the "weights" or quantities.

Such a general revision was made beginning with the first week of January, 1924.

When my first calculations were made (in 1922) the Census figures for 1919 had not yet been utilized by the United States Bureau of Labor Statistics for revising its own index number, so that I began with the "weights" (quantities produced) used by the War Industries Board. These were based on statistics of production of 1917, unfortunately a war year.

For the 1924 index I have changed from the system of quantities produced in 1917 to that of quantities sold in 1919 (based on the figures now used by the Bureau of Labor Statistics). The actual influence of this change, however, is very small.[1]

Of more practical importance is the addition to the list of commodities of coal (both anthracite and bituminous) and cotton.

The addition of coal was made possible by the fact that Dun's Review (from which almost all of my quotations are taken) added this important item, or series of items, to its list of weekly quotations. Spot cotton quotations are taken from the daily papers and averaged through each week ending Thursday (Friday's quotation not being available in time).

Several other, but minor, revisions were also made.

Table I gives both the old index (*i. e.*, for 1923) and the revised (*i. e.*, beginning with 1924).

As explained in last year's article, strictly speaking the *base* for the

[1] The not uncommon impression that a change in weights makes a material difference in the index figure is erroneous. It has been frequently stated, for instance, that the big difference between the revised index of the United States Bureau of Labor Statistics and their old index (amounting in one case to 10 per cent) was due to a change in weights from those based on 1909 figures to those based on 1919 figures. It turns out, however, that this change in weights made an almost negligible difference and that the major part of the discrepancy is accounted for by the introduction of new commodities. See *Labor Review*, May, 1924, p. 69.

1923 series is not 1913 but the mid-November week (ending November 17) of 1922, and the corresponding *base number* was not 100 but 156, that figure being taken because it was the index number of the United States Bureau of Labor Statistics for the month of November, 1922.

In the same way the base for the 1924 series is the mid-November week of 1923 (ending November 16) and the base number was 151.9, that being the index number for that week by the first, or 1923, series.

Thus, Table I really involves three series: that of the United States

TABLE I

WEEKLY INDEX NUMBERS

For week ending Friday	Index number of prices in per cent of pre-war, *i. e.*, 1913 = 100 per cent	Purchasing power of the dollar in pre-war cents, *i. e.*, 1913 = 100 cents
1923		
January 5	155.9	64.1
January 12	156.0	64.1
January 19	156.8	63.8
January 26	157.5	63.5
February 2	158.7	63.0
February 9	158.4	63.1
February 16	161.3	62.0
February 23	163.9	61.0
March 2	164.8	60.7
March 9	164.8	60.7
March 16	166.1	60.2
March 23	166.6	60.0
March 30	167.2	59.8
April 6	168.0	59.5
April 13	167.1	59.9
April 20	166.8	60.0
April 27	166.3	60.1
May 4	164.8	60.7
May 11	163.7	61.1
May 18	162.7	61.5
May 25	161.6	61.9
June 1	159.8	62.6
June 8	160.3	62.4
June 15	159.3	62.8
June 22	158.4	63.1
June 29	155.0	64.5
July 6	154.3	64.8
July 13	153.2	65.3
July 20	153.2	65.3
July 27	153.7	65.1
August 3	153.0	65.4
August 10	152.9	65.4
August 17	153.2	65.3
August 24	154.6	64.7
August 31	154.8	64.6
September 7	154.9	64.6
September 14	155.6	64.3
September 21	156.0	64.1
September 28	154.9	64.6
October 5	156.7	63.8
October 12	156.6	63.9
October 19	155.6	64.3
October 26	155.0	64.5
November 2	153.5	65.4
November 9	154.2	64.8
November 16	151.9	65.9
November 23	151.4	66.1
November 30	150.7	66.4
December 7	151.1	66.2
December 14	150.0	66.7
December 21	150.4	66.5
December 28	150.9	66.3

TABLE I (Continued)

WEEKLY INDEX NUMBERS

For week ending Friday	Index number of prices in per cent of pre-war, i. e., 1913 = 100 per cent	Purchasing power of the dollar in pre-war cents, i. e., 1913 = 100 cents
1924		
January 4	150.0	66.7
January 11	150.0	66.7
January 18	149.8	66.8
January 25	150.1	66.6
February 1	152.5	65.6
February 8	154.2	64.8
February 15	153.2	65.3
February 22	151.5	66.0
February 29	150.9	66.3
March 7	150.8	66.3
March 14	149.7	66.8
March 21	148.2	67.5
March 28	146.0	68.5
April 4	146.2	68.4
April 11	146.6	68.2
April 18	145.6	68.7
April 25	145.8	68.6
May 2	145.2	68.9
May 9	145.0	69.0
May 16	145.2	68.9
May 23	144.8	69.1
May 30	144.5	69.2
June 6	143.3	69.8
June 13	143.4	69.8
June 20	142.7	70.1
June 27	143.1	69.9
July 4	143.2	69.9
July 11	142.3	70.3
July 18	144.5	69.2
July 25	147.1	68.0
Aug. 1	148.7	67.2
Aug. 8	147.9	67.6
Aug. 15	148.5	67.4
Aug. 22	147.1	68.0

Bureau of Labor Statistics (for November 1922—relatively to 1913=100), that of the first (1923) weekly index, and that of the second (1924) index. These three series are outlined in Table II, where the base number of each series is given in dark faced type:

TABLE II

THE THREE INDEX NUMBERS LINKED TOGETHER

	U. S. Bureau of Labor statistics	Weekly index 1923	Weekly index 1924
1913	**100**		
November, 1922	156 =		
November, 1923		**156**	
January (first week) 1924		151.9 =	**151.9**
			150.0

While, theoretically, the three indexes are three different series, the discrepancies are not great. It is impossible to make any complete comparisons between the three series, because of lack of data. But we

may approximately compare them as follows: Out of the 200 commodities used in my 1923 index number (all from Dun's) we find only 157 available for the week ending January 17, 1913. This, the mid-week of January, may be taken as typical for the month of January, and the month of January, in turn, may be taken as typical for the year 1913 (because the Bureau of Labor Statistics index number for January, 1913, is the same as for the entire year 1913, *i. e.*, 100). The results are given in Table III (in which, as in Table II, the base numbers are dark faced).

TABLE III

TEST COMPARISON BETWEEN UNITED STATES BUREAU OF LABOR STATISTICS AND 1923 WEEKLY INDEX NUMBER

	U. S. B. L. S.	Weekly index number 1923 series (157 commodities)
1913	**100**	96.1
November, 1922	156 =	**156.**

That is, the discrepancy (between 100 and 96.1) is only 4 per cent although the two dates are nearly ten years apart.

Similarly we may test the two systems of weekly indexes, those for 1923 and 1924, as shown in Table IV (base numbers, dark faced).

TABLE IV

TEST COMPARISON BETWEEN THE 1923 AND 1924 SYSTEMS OF WEEKLY INDEXES

	1923 system	1924 system
Week ending, November 17, 1922	**156**	151.8
Week ending, November 16, 1923	151.9 =	**151.9**
Week ending, January 4, 1924	151.5	150.0

That is, for January 1924, the discrepancy (between 151.5 and 150.0) is 1 per cent, while for November, 1922, the discrepancy (between 156 and 151.8) is 2.7 per cent.

It is the inclusion of cotton and coal which makes the chief difference. Thus, if in the 1924 system cotton and coal be omitted, the 151.8 becomes 154.9 which differs from 156 by only .7 per cent. In the same way the 1924 method—omitting cotton and coal—gives, for January 4, 1924, not 150.0 but 151.7, so that the discrepancy is only between 151.5 and 151.7 or .13 per cent.[1]

[1] The agreement elsewhere may be greater or less. The Bureau's index changes more slowly than mine, both upward and downward, so that, at the maximum and minimum points, we may expect the

Thus, to all intents and purposes, the three series really form three links in one chain, each link beginning practically where the preceding one ends. The only week-to-week disturbance in Table I, caused by linking the 1924 system to the 1923 system, is a slight downward dislocation between the last week of 1923 and the first week of 1924. Had either the 1923 or the 1924 system been used for both these weeks the index would have risen slightly instead of fallen. (Thus, under the 1923 system the change would have been from 150.9 to 151.5, instead of to 150.0).

Table V gives the complete list of commodities under the revised system, with their "weights" (quantities) and, for comparison, the "weights" used in the previous year. The commodities dropped out this year (twenty in number but mostly of slight importance) are not included in the table. The commodities are listed in the order of their importance as indicated by their values (price times quantities) in mid-November, 1923:

Evidently the index number may simply be described as giving the relative value, from week to week, of a cargo of the 205 commodities in the above specified quantities.

Of these 205 commodities, 118 have the same specified quantities, or "weights," in 1924 as in 1923 and many of the others have not changed greatly.

It will be seen that the most important commodity is cotton, which has a value of 1357 ($1,357,000,000). This is 7 per cent of the total ($18,745,230,000). Beef comes next with 6.8 per cent. The first 14 commodities constitute 50 per cent of the total.

Round weights are used, as last year, when a commodity is unimportant and also when the "true" weight was sufficiently close to the round one used.

In the revised index number it has been found possible, by adjusting the number of commodities in the various groups, to dispense with last year's "class weights" for all groups with the exception of the chemical group. For this group a class weight of 1.5 is used this year. That is, the "weight" (quantity) of every commodity used in this group was increased 50 per cent, the reason being that it was impossible to find

---

greatest differences. The maximum of the Bureau's index was for May 1920 when it stood at 247. Computing my own index for May 1920 for the 159 commodities available, I find it to be 274.2 (relatively to 156 for mid-November 1922) which is 11 per cent higher than the Bureau's figure (also considered relatively to 156 for November 1922); in fact, relatively to 1913, mine exceeds the Bureau's by even more than 11 per cent (namely 15 per cent) since mine for 1913 is 96.1 as against the 100 for the Bureau. (It is interesting to note that the Bureau's own original index for May 1920 was 272 before its recent revision.) When the minimum was reached by the Bureau's index in January 1922 its figure stood at 138 as against 133.2 by my method for 199 commodities, mine being lower than the Bureau's by 3.8 per cent. (By the Bureau's old system the figure for January 1922 was 148.)

TABLE V

LIST OF COMMODITIES IN THE ORDER OF THEIR IMPORTANCE*

	Commodity †	Unit	Number of units (000,000 omitted)		Value (000,000 omitted)
			Revised 1924	Used in 1923	
(Cl)†	Cotton, raw, New York	lb.	4,000.	not used	1,357.
(Fo)	Beef, steers, live	100 lbs.	130.	122.	1,300.
(Fu)	Petroleum, crude, at well	bbl.	400.	365.	940.
(Fu)	Coal, Bituminous Pool 1 (N. S.)	ton	300.	not used	938.
(Fo)	Wheat, No. 2 R	bu.	700.	603.	861.
(Fo)	Sugar, fine gran. in bbls	100 lbs.	80.	40.2	700.
(Fo)	Hogs, live	" "	100.	104.	665.
(Fo)	Eggs, fresh gathered firsts	doz.	1,000.	100.	520.
(Fo)	Butter, state dairy, tubs, finest	lb.	1,000.	not used	515.
(Fo)	Sugar, Cent. 96°	100 lbs.	70.	66.2	510.
(Fo)	Butter, creamery, extra	lb.	700.	671.	371.
(Fo)	Flour, spring patents	196 lbs.	60.	53.4	360.
(Fu)	Coal, Anth., chestnut (Indpt.)	ton	30.	not used	345.
(Fo)	Corn No. 2 yellow	bu.	300.	1,260.	325.
(Fo)	Hay No. 1	100 lbs.	200.	360.	300.
(Fo)	Flour, winter soft straights	196 "	60.	53.4	285.
(M)	Tank plates, Pittsburgh	100 "	100.	same	250.
(M)	Steel bars, Pittsburgh	" "	100.	"	240.
(Fo)	Tobacco, L'vlle, medium	lb.	1,000.	100.	240.
(Cl)	Silk, Japan, Fil. No. 1, Sinshiu	"	30.	38.	234.
(Fo)	Pork, mess	bbl.	10.	9.	230.
(M)	Billets, open hearth, Philadelphia	ton	5.	27.	226.
(M)	Sheets, black No. 28, Pittsburgh	100 lbs.	50.	34.4	188.
(Cl)	Standard clay worsted, 16 oz	yd.	50.	same	181.
(B)	Brick, Hud. River, common	1,000	10.	"	180.
(B)	Portland Cement, Northampton, Pa. mill	bbl.	90.	92.8	167.
(Fo)	Hams, N. Y. big in tcs	lb.	1,000.	same	163.
(Fu)	Coal, Bituminous, Pool 34 (High Vol. St.)	ton	100.	not used	163.
(P)	Paper, news roll	100 lbs.	40.	40.1	160.
(Fu)	Gasolene, auto in gar. st. bbls	gal.	1,000.	100.	155.
(Fo)	Lard, N. Y. Mid. W	100 lbs.	10.	same	148.
(Fo)	Coffee, Santos, No. 4	lb.	1,000.	"	145.
(Fu)	Kerosene, wagon, delvd	gal.	1,000.	100.	140.
(M)	Copper, electrolytic	lb.	1,000.	same	134.
(Cl)	Wool, territory, scoured basis, fine staple choice	"			
		lb.	100.	10.	128.
(Cl)	Wool, pulled, Delaine	"	100.	not used	120.
(B)	F. A. S.‡ Pl. Wh. Oak 4/4"	M ft.	1.	same	120.
(Fo)	Bacon, N. Y. 140s down	lb.	1,000.	"	120.
(Fo)	Cottonseed oil	"	1,000.	"	119.
(Fu)	Coal, Anthracite, stove (Indpt.)	ton	10.	not used	115.
(Cl)	Wool, territory, scoured basis, half-blood combing	lb.	100.	10.	115.
(Cl)	Print cloths, 38 1/2 in. 64 x 60	yd.	1,000.	not used	109.
(B)	F. A. S.‡ H. Maple 4/4"	M ft.	1.	same	105.
(M)	Pig iron, basic valley furnace	ton	5.	14.3	103.
(Cl)	Wool, Calif., scoured basis, southern	lb.	100.	not used	100.
(B)	Linseed oil, city raw	gal.	100.	same	96.
(Cl)	Staple ginghams	yd.	500.	1,000.	95.
(Fo)	Sheep, live	100 lbs.	10.	same	95.
(Cl)	Silk, China, St. Fil 1st	lb.	10.	"	93.
(Fo)	Cod, Grand Banks	100 lbs.	10.	"	90.
(Fo)	Apples, common	bbl.	30.	52.4	90.
(Cl)	Brown sheetings, 4 yd. per lb	yd.	600.	not used	84.
(Fo)	Eggs, nearby fancy	doz.	100.	same	84.
(Ch)	Cocoanut oil, crude tks., f. o. b. coast	lb.	1,000.	100.	83.
(Fu)	Coal, Anthracite, pea (Indpt.)	ton	10.	not used	80.
(Cl)	Hides, country No. 1 steers	lb.	1,000.	100.	80.
(M)	Pig iron, Bessemer, Pittsburgh	ton	3.	11.9	79.
(Fo)	Rye, No. 2	bu.	100.	same	77.
(P)	Burlap, 10 1/2 oz., 40 in	yd.	10.	1,000.	77.
(P)	Paper, book, S. S. & C	lb.	1,000.	same	73.
(M)	Lead, N. Y	"	1,000.	"	69.
(M)	Spelter, N. Y	"	1,000.	not used	68.
(Fo)	Beans, pea choice	100 lbs.	10.	1.	68.
(Cl)	Leather, Belting butts, No. 1, light	lb.	100.	same	62.
(P)	Paper, boards, chip	ton	1.	"	60.
(Ch)	Rosin "B"	bbl.	10.	"	58.
(Cl)	Bleached sheetings, st	yd.	300.	1,000.	57.
(B)	Long leaf yellow pine timbers, 12 x 12"	M ft.	1.	same	56.
(Cl)	Brown drills, standard	yd.	300.	1,000.	56.
(Ch)	Acid, sulphuric, 60°	100 lbs.	100.	same	55.

TABLE V (*Continued*)

	Commodity †	Unit	Number of units (000,000 omitted)		Value (000,000 omitted)
			Revised 1924	Used in 1923	
(M)	Tinplate, Pittsburgh	100-lb. box	10.	same	55.
(B)	No. 1 Com. yellow pine boards 1 x 4"	M ft.	1.	"	51.
(M)	Wire rods, Pittsburgh	ton	1.	"	51.
(Fo)	Oats, No. 3, white	bu.	100.	"	51.
(Ch)	Alcohol, denatured, form 5	gal.	100.	"	51.
(M)	Galvanized sheets No. 28 Pittsburgh	100 lbs.	10.	"	50.
(Ch)	Alcohol, 190 prf. U. S. P.	gal.	10.	not used	48.
(Fu)	Coke, Conn'ville oven, foundry, prompt ship.	ton	10.	45.5	48.
(Cl)	Broadcloth, 54 in.	yd.	10.	not used	46.
(M)	Open hearth rails, hy. at mill	ton	1.	same	43.
(M)	Tin, N. Y.	lb.	100.	"	43.
(Cl)	Serge, 16 oz.	yd.	10.	50.	42.
(Cl)	Leather, scoured oak backs, No. 1	lb.	100.	same	42.
(M)	Billets, Bessemer, Pittsburgh	ton	1.	8.96	40.
(B)	Penn. hemlock, b. price	M ft.	1.	same	40.
(M)	Barb wire, galvanized, Pittsburgh	100 lbs.	10.	"	38.
(Cl)	Leather, Union backs, t. r. l. b.	lb.	100.	"	38.
(Fu)	Coke, Conn'ville oven, furnace, prompt ship.	ton	10.	56.7	38.
(B)	Douglas fir timbers, 12 x 12"	M ft.	1.	same	37.
(Ch)	Alum, lump	lb.	1,000.	"	35.
(Fo)	Rice, foreign, Saigon, No. 1	lb.	1,000.	"	34.
(Ch)	Caustic soda 76%	100 lbs.	10.	same	32.
(Fo)	Salt, 280 lb. bbl.	bbl.	10.	1.	32.
(Ch)	Sulphate ammonia, dom., f. o. b. works	100 lbs.	10.	same	30.
(Cl)	Serge, 11 oz.	yd.	10.	50.	29.
(Cl)	36-in. cotton-warp serge	"	50.	100.	29.
(P)	Rubber, Plan. 1st Latex, crude	lb.	100.	same	29.
(Cl)	Fancy cassimere, 13 oz.	yd.	10.	100.	27.
(M)	Iron bars, refined, Philadelphia	100 lbs.	10.	same	27.
(Fo)	Cheese, N. Y. fl. held sp.	lb.	100.	not used	26.
(Ch)	Acid, carbolic, drums	"	100.	" "	26.
(M)	Aluminum, pig (ton lots)	"	100.	same	26.
(Fo)	Onions	bag	10.	"	25.
(Ch)	Nitrate soda	100 lbs.	10.	"	24.
(P)	Rubber, up river fine	lb.	100.	"	24.
(M)	Pig iron, No. 2, So. Cincinnati	ton	1.	4.72	23.
(Ch)	Bicarbonate soda, Am.	100 lbs.	10.	same	23.
(B)	China wood oil (bbls. spot)	lb.	100.	"	22.
(Ch)	Castile soap, pure white	"	100.	"	22.
(B)	Lime, f. o. b. fty.	200 lb. bbl.	10.	100.	19.
(B)	No. 1. Com. mahogany 4/4"	M ft.	.1	not used	18.
(Fo)	Tobacco, L'vlle, common	lb.	100.	same	18.
(Ch)	Brimstone, crude, dom.	ton	1.	"	18.
(B)	Qtd. Oak, AA Grade G. I S.§	M ft.	.1	not used	17.
(Ch)	Glycerine, C. P. in bulk	lb.	100.	same	17.
(B)	F. A. S.‡ Birch (red) 4/4"	M ft.	.1	"	15.
(Ch)	Soda Ash 58% light	100 lbs.	10.	100.	14.
(Fo)	Tobacco, L'vlle Burley red—Com. sht.	lb.	100.	1,000.	14.
(B)	F. A. S.‡ Poplar 4/4"	M ft.	.1	same	13.
(B)	F. A. S.‡ Chestnut 4/4"	" "	.1	"	13.
(B)	F. A. S.‡ ash 4/4"	" "	.1	"	12.
(M)	Platinum	oz.	.1	"	12.
(Fo)	Corn oil	lb.	100.	"	12.
(Ch)	Soya bean oil tk. coast, spot	"	100.	"	11.
(Fo)	Coffee, No. 7 Rio	"	100.	"	11.
(Fo)	Apples, evaporated choice	"	100.	10.	11.
(B)	Red Lead, American	"	100.	same	11.
(P)	Hemp, midway ship	"	100.	"	11.
(Fo)	Prunes, Cal. 40–50, 25 lb. box	"	100.	"	11.
(B)	F. A. S.‡ Cypress 4/4" (old grades)	M ft.	.1	"	10.
(B)	Litharge, Am.	lb.	100.	"	10.
(Fo)	Cabbage	bbl.	10.	"	10.
(B)	Zinc, F. P. R. S. (paints)	lb.	100.	"	10.
(Ch)	Turpentine	gal.	10.	100	10.
(Ch)	Indigo Madras	lb.	10.	same	10.
(Fo)	Raisins, California stand., loose mus.	"	100.	"	10.
(B)	F. A. S.‡ Pl. red gum 4/4"	M ft.	.1	same	10.
(Ch)	Camphor, domestic	lb.	10.	"	9.
(Ch)	Alcohol, wood, 95 p. c.	gal.	10.	not used	9.
(Ch)	Acid, muriatic, 18°	100 lbs.	10.	" "	9.
(Fo)	Peaches, California standard	lb.	100.	same	9.
(B)	Shellac, D. C.	"	10.	"	9.
(Ch)	Opium, jobbing lots	"	1.	.1	8.

TABLE V (*Continued*)

	Commodity †	Unit	Number of units (000,000 omitted)		Value (000,000 omitted)
			Revised 1924	Used in 1923	
(B)	Zinc, American (paints)	lb.	100.	same	8.
(Fo)	Tallow, N. Y. sp. loose	"	100.	"	8.
(Ch)	Palm oil, Lagos	"	100.	"	8.
(Fo)	Peas, Scotch, choice	100 lbs.	1.	"	8.
(Cl)	36 in. all worsted serge	yd.	10.	100.	7.
(Cl)	36 in. all worsted Panama	"	10.	100.	7.
(Ch)	Vitriol, blue	lb.	100.	not used	7.
(Ch)	Morphine sulphate (bulk)	oz.	1.	same	6.
(M)	Quicksilver, 75 lb. flask	flask	.1	"	6.
(Fo)	Molasses, extra fancy	gal.	10.	"	6.
(Ch)	Borax, crystal in bbls.	lb.	100.	"	6.
(Ch)	Acid, nitric, 42°	lb.	100.	1,000.	5.
(Ch)	Quinine, 100 oz. tins	oz.	10.	same	5.
(Ch)	Acid, citric, domestic	lb.	10.	"	5.
(B)	Adirondack spruce 2 x 4"	M ft.	.1	"	5.
(Fo)	Tobacco, L'ville fine	lb.	10.	"	4.
(Fo)	Tea, Hyson, firsts	"	10.	"	4.
(Ch)	Acid, acetic, 28 deg.	100 lbs.	1.	"	3.
(Ch)	Tartaric crystals	lb.	10.	"	3.
(Fo)	Tea, Japan, low	"	10.	"	3.
(Ch)	Muriate potash, 80% unit‖		.1	not used	3.
(Fo)	Syrup, sugar, medium	gal.	10.	same	3.
(Fo)	Tea, Formosa, fine	lb.	10.	"	3.
(Ch)	Cream tartar 99%	"	10.	.1	3.
(Ch)	Bones, ground, steamed, etc.	ton	.1	same	2.
(Ch)	Epsom salts	100 lbs.	1.	not used	2.
(Fo)	Tea, Formosa, fair	lb.	10.	same	2.
(Fo)	Tea, Hyson, low	"	10.	"	2.
(Fo)	Molasses, blackstrap	gal.	10.	"	2.
(Ch)	Licorice root	lb.	10.	"	2.
(Ch)	Bleaching powder, over 34%	100 lbs.	1.	10.	2.
(Ch)	Menthol, cases	lb.	.1	1.	2.
(Ch)	Arsenic, white	"	10.	100.	1.
(Ch)	Castor oil, No. 1	"	10.	same	1.
(Ch)	Sal soda, American	100 lbs.	1.	not used	1.
(Ch)	Acid, stearic single pressed	lb.	10.	" "	1.
(Ch)	Calomel, American	"	1.	same	1.
(B)	White lead, in oil	"	10.	"	1.
(Cl)	Hides No. 1, calfskin	"	10.	"	1.
(Fo)	Pepper, Singapore, black	"	10.	"	1.
(Ch)	Formaldehyde	"	10.	100.	1.
(Cl)	Hides, No. 1 kip	"	10.	same	1.
(Cl)	Hides, No. 1 buff hides	"	10.	100.	.9
(B)	White lead, dry	"	10.	same	.9
(Ch)	Nux vomica, powdered	"	10.	"	.9
(Ch)	Gambier	"	10.	"	.9
(M)	Antimony, ordinary	lb.	10.	same	.9
(Ch)	Chlorate potash	"	10.	100.	.7
(Ch)	Cocaine, Hydrochloride	oz.	.1	same	.7
(Cl)	Hides, No. 1 extremes	lb.	10.	100.	.7
(Fo)	Mace	"	1.	same	.6
(Fo)	Tea, Japan, best	"	1.	"	.5
(Ch)	Nitrate silver, crystals	oz.	1.	"	.4
(Ch)	Chloroform	lb.	1.	10.	.4
(Fo)	Cloves, Zansibar	"	1.	same	.4
(Fo)	Nutmegs 105s–110s	"	1.	"	.3
(Ch)	Indigo paste, 20%	"	1.	10.	.3
(B)	Ochre, French	"	10.	same	.2
(Fo)	Ginger, Cochin	"	1.	"	.2
(Ch)	Balsam, Peru	"	.1	"	.2
(B)	Whiting, commercial	100 lbs.	.1	"	.1
(Ch)	Gum arabic, benzoin, Sumatra	lb.	.1	1.	.03
					18,745.23

* As measured by value (last column), *i. e.*, product of quantity and price (in second week of November, 1923),—the revised value used for 1924.

†The groups (as summarised later in Table VI) are indicated at the left as follows: Fo =foods; Cl =cloths and clothing; P =paper, rubber, fiber; M =metals; Fu =fuels; B =building material; Ch =chemicals.

‡F. A. S. is an abbreviation for Firsts and Seconds.

§G. 1 S. is an abbreviation for Good on One Side.

‖80% unit means 80% pure.

enough commodities in this group to give it its proper representation in relation to the other groups except by thus letting every two chemicals stand for three.

Table VI shows the comparative importance of the various classes of commodities both in the revised index and in last year's index.

TABLE VI

ASSORTMENT, OR RELATIVE IMPORTANCE, OF GROUPS OF COMMODITIES, IN THE WEEKLY INDEX

Group	Per cent of total value	
	1923	1924
Food	50.5	45.9
Clothing and Cloths	16.8	16.9
Paper, Rubber, Fibers	2.5	2.3
Metals	10.0	9.5
Fuels	10.3	15.9
Building Materials	6.0	5.9
Chemicals	3.9	3.6
	100.0	100.0

The only important changes in these percentages are the relative increase in the representation of fuels and decrease in that of foods. Both these changes and the specific additions and substitutions of commodities afford, I believe, some improvement in the index number.

Since my weekly index has been published, Norman Crump of the *London Financial Times* has published a weekly index for England. His index and mine are now exchanged by cable and published in English and American papers. Crump's index includes 73 commodities. The price ratios are averaged geometrically.

Both indexes usually anticipate the monthly indexes and, for this reason alone, have a practical value, quite aside from the fact that they are always more nearly "up to the minute."

The three charts explain themselves. Chart I compares the only three weekly index numbers published one (Gibson's) being for foods only.

Chart II compares my index *nearest the beginning of each month* with Bradstreet's and Dun's for the first of each month.

Chart III compares the *monthly average* of my index with those of the United States Bureau of Labor Statistics.

There is a family resemblance between my index number and each of the others, even Crump's for England. The greatest similarity is with the United States Bureau of Labor Statistics.

The comparisons show that, relatively to my index:

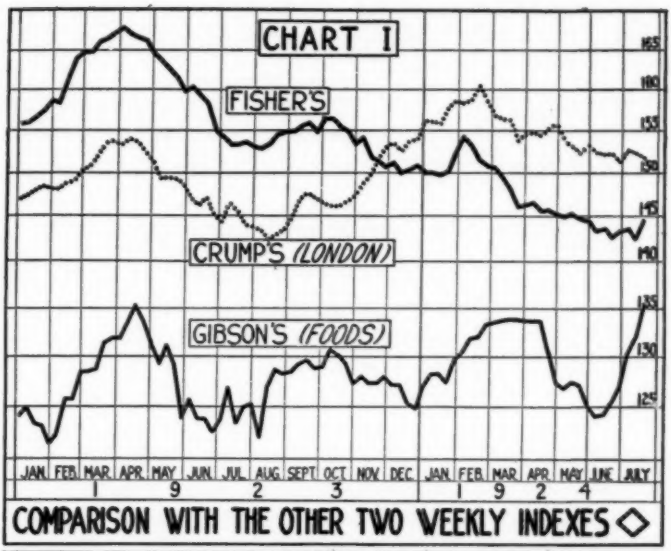


CHART I
FISHER'S
CRUMP'S (LONDON)
GIBSON'S (FOODS)
JAN FEB MAR APR MAY JUN JUL AUG SEPT OCT NOV DEC JAN FEB MAR APR MAY JUNE JULY
1 9 2 3 1 9 2 4
COMPARISON WITH THE OTHER TWO WEEKLY INDEXES ◇

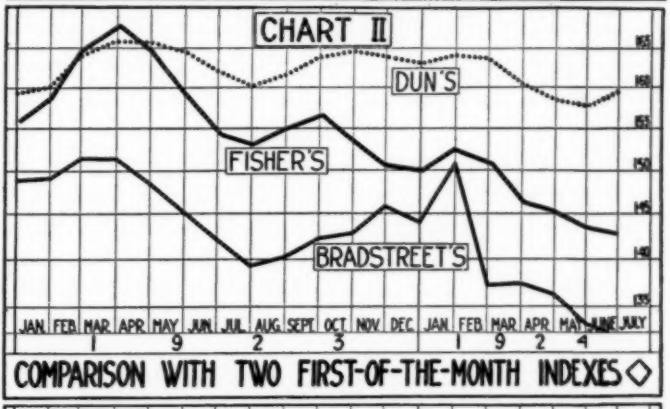


CHART II
DUN'S
FISHER'S
BRADSTREET'S
JAN FEB MAR APR MAY JUN JUL AUG SEPT OCT NOV DEC JAN FEB MAR APR MAY JUNE JULY
1 9 2 3 1 9 2 4
COMPARISON WITH TWO FIRST-OF-THE-MONTH INDEXES ◇

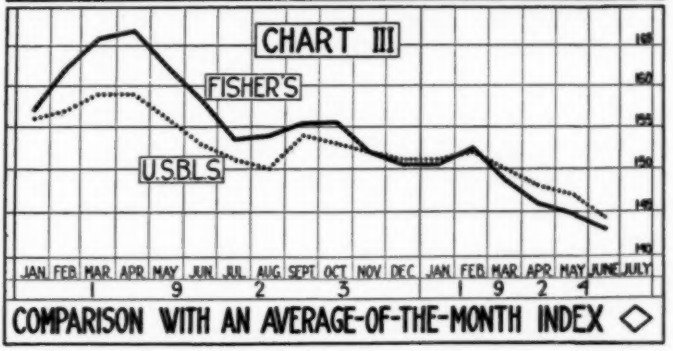


CHART III
FISHER'S
U.S.B.L.S.
JAN FEB MAR APR MAY JUN JUL AUG SEPT OCT NOV DEC JAN FEB MAR APR MAY JUNE JULY
1 9 2 3 1 9 2 4
COMPARISON WITH AN AVERAGE-OF-THE-MONTH INDEX ◇

The Crump index varies from 7 per cent above to 9 per cent below.

The Gibson index (for foods only) varies from 6 per cent below to 23 per cent below.

Bradstreet's index varies from 1 per cent below to 10 per cent below.

Dun's index varies from 11 per cent above to 1 per cent below.

The Labor Bureau's index varies from 2 per cent above to 5 per cent below.

The *average* deviations of these five indexes from mine are:

Crump, −2.2 per cent
Gibson, −16.4 per cent
Bradstreet, −6.9 per cent
Dun, +5.6 per cent
Labor Bureau, −1.2 per cent

While these averages (and most of the weekly and monthly discrepancies as well) are not inordinately large, they are larger than are commonly thought of between homogeneous indexes.

The reasons for their being so great are:

1. The first three indexes are *not* homogeneous with mine.

Crump's index is for a market three thousand miles away, and one in which the prices are expressed in paper money, instead of in gold (to say nothing of the less important causes for deviations, namely the smaller number of commodities used, the difference in assortment, and in formula employed in the calculation). Despite these facts, there is a surprising resemblance between the week to week movements of the two indexes.

Gibson's index is for foods only while only 46 per cent of my index data consists of foods. The prices of foods and farm products were dislocated by the sudden alternation of war and peace. This fact is, of course, at the bottom of the discontent of the farmer. It is interesting to note that the discrepancy between the food index and my general index has tended gradually to disappear, having changed from 20 per cent in January 1923 to 6 per cent in July 1924.

Bradstreet's index is more nearly homogeneous with mine than the two preceding, but it has "haphazard weighting" instead of a weighting based on the statistical importance of the various commodities.[1]

2. The common (and correct) impression of a closer agreement between different indexes than that here found is based on comparing *annual* indexes. Naturally the monthly and weekly fluctuations are considerably greater.

3. The years 1923 and 1924 are abnormal years, in which readjust-

[1] See my *Making of Index Numbers*, p. 333.

ments from the war and the crisis of 1920 are still taking place so that there is a greater scattering of prices than usual and consequently a greater opportunity for discrepancies between different indexes.

4. As to the Labor Bureau's index, it shows less change than my index number, in both upward and downward movements. This relative invariability of the Bureau's index seems to be due principally to two facts: (1) the lesser emphasis on metals which are apt to change sooner and further than the rest, and (2) the greater number of insensitive commodities (such as matches), the prices of which vary seldom and little. It would seem that, for most practical purposes, the inclusion of very insensitive commodities reduces, rather than increases, the usefulness of an index number.

THE RELATION OF THE UNITED STATES TO INTERNATIONAL STATISTICS

By Walter F. Willcox, *Cornell University*

European statists with occasional and scanty help from those in the United States and other non-European countries have tried for more than seventy years to build up a system of international coöperation in statistics and have made some progress. A brief record of their work, their successes and their failures, may be of interest to members of the American Statistical Association and prove helpful later in case statists from abroad should come again across the Atlantic as guests of the country and the society.

The history of international statistics, like nearly all forms of history, clusters about organizations or institutions. In this formal sense it began with the London Exhibition of 1851 and the International Statistical Congress at Brussels in 1853. It is my purpose to sketch first the development of international statistics and then the relation of the United States government and of American scientific organizations to that development.

I

Towards the middle of the nineteenth century the spirit of internationalism seems to have been in the air and perhaps to have been strongest in England. The British adoption of free trade and the belief that a similar change of policy would quickly come in other countries were doubtless parts of the general trend. These conditions prepared the way for a movement headed by the Prince Consort to organize an international exhibition in London. Within the preceding century national exhibitions of limited range had been held in England and had spread to France, where indeed a proposal was made to extend the scope of a French national exhibition and make it international. The French Minister of Commerce frowned on the proposal, condemning it as probably emanating from enemies of French industry. In England several of these national exhibitions were conducted by the Society for the Encouragement of Arts, Commerce and Manufactures, which in the late '40's was reorganized under the name of the Society of Arts and chose for president the Prince Consort. After some preliminary steps he called a meeting of the society at Buckingham Palace and laid before them his proposal to organize an international exhibi-

tion. The plan encountered opposition at first, but the tact, energy and enthusiasm of the Prince Consort triumphed over all obstacles and the exhibition was carried to a success which was quite unforeseen and triumphant.

Some of the enthusiastic advocates of the innovation, notably Baron Bunsen, had included in their dream the hope that the London Exhibition would bring together scholars and scientists and thus facilitate international intercourse of another sort. At that period Quetelet was the best known international statist; he had been a teacher of the Prince Consort; he had suggested and been largely responsible for the organization of the Statistical Society of London in 1834 and had revolutionized European census practice by the Belgian census of 1846. He attended the London Exhibition in June 1851,[1] conferred with a number of British statists and on his return to Brussels joined Visschers in proposing that soon after the close of the London Exhibition an international statistical congress should be held for the purpose of securing greater comparability in statistical publications. The plan was provisionally approved and Visschers on a subsequent trip to London was authorized to inquire into the attitude of the English statists. He reported that they would cordially support the plan, but were in favor of leaving the initiative to the Belgians and of having the meeting at Brussels. The part of the United States in these negotiations will appear later. Here it is enough to note that the congress achieved a notable success. Many of the leading statists of Europe attended, formed or ripened acquaintance with one another and discussed their common problems. This congress was the first of a series of nine held at intervals of two or three years and all in European cities. But after the congresses became better known and it was realized that any one who chose to take out membership could attend and participate, the relatively small number of technical experts were more and more lost in the increasing body of amateurs or interested bystanders. To remedy this growing defect Quetelet proposed at the fourth congress in London that the official delegates or technical experts should meet by themselves immediately before or after the main congress. At the time nothing came of his proposal, but at the succeeding congress it was taken up by Engel with more energy and system.

Quetelet, repeating his London suggestion, had asked for a preliminary meeting of the official delegates to the Berlin Congress at which they might confer about questions of professional interest and

[1] The doubt expressed by Lottin (*Quetelet, Statisticien et Sociologue*, p. 81, footnote) that Quetelet went to London at this time is apparently set at rest by the mention of his attending a meeting of the Statistical Society of London in June of that year (Stat. Soc. of London, Journal, Vol. XIV, p. 261).

the German preparatory committee had devoted to that purpose the Friday and Saturday preceding the week of the congress. Quetelet and Engel, however, differed about the object of this preliminary meeting. Quetelet wished to organize work in international statistics on a coöperative basis. He had already made good progress with a report which was published two years later and dealt with international population statistics. He hoped that the official statists of other countries would make similar reports in other fields, a hope which subsequently was realized. So he desired to limit participation in the conference of experts to those who were sharing in this coöperative effort. But Engel believed that, so long as the program of each congress was determined and the preliminary studies made by a committee of statists from the country which acted as host, the congresses could not become really international. He argued that if the congresses were to form a connected series and to deserve their title of international the preparations for each meeting must be in the hands of a permanent international committee on which the point of view of every important country would be represented.

With this proposal of Engel it is suggestive to compare the action taken fifty years later by the Second Hague Peace Conference. That body pointed out the need of preparing long in advance the program for the third conference and voted: "That some two years before the probable date of the meeting a preparatory committee should be charged . . . with the task of collecting the various proposals to be submitted to the conference, of ascertaining what subjects are ripe for embodiment in an international regulation and of preparing a program."

Engel's plan for a permanent international committee, meant a wide departure from the preceding practice. At the Berlin meeting it was finally referred to a special committee. At later congresses the idea of Quetelet gained support. At the seventh it ripened into a plan for a great coöperative work on international statistics, the Statistique Internationale de l'Europe. This led to the creation by the eighth congress of a permanent commission. It did not confine its sessions to the times and places of congress meetings but met twice between the eighth and the ninth sessions. It held its last meeting at Paris in 1878 and devoted itself mainly to perfecting its own organization and enlarging its field of work. It proposed to begin the publication in French and at Paris of an international statistical annual and bulletin. Upon these resolutions the German representatives probably acting on a hint from Bismarck declined to commit themselves and when plans for a meeting of the commission at Rome in 1879 were well advanced a

number of countries opposed accepting either that invitation or another for the tenth congress to be held in 1880. As the decision to give the permanent commission a home in Paris was distasteful to Bismarck he forbade the Prussian statists to attend further meetings. Other German states followed the lead of Prussia in this matter[1] and in consequence the president of the permanent commission, Keleti of Hungary, a month before the session at Rome was to have been held was obliged to write to the other members: "All the German states have refused to be represented on the permanent commission; Switzerland has followed this example; representatives of Spain and of Sweden have sent excuses; Portugal and Holland have not answered; regarding England and Russia I am not even yet in a position to know whether they will be officially represented or not. Under these circumstances . . . I have at last decided to postpone the session."[2] And thus the series of meetings both of congress and of commission came to an end.

The main shortcomings of these gatherings were:

1. A majority of those who attended came from the city in which the congress was held or from its immediate vicinity and were attracted by unprofessional motives. Thus the announcement at Vienna in 1857 that the members would be given by the Emperor a free Sunday excursion induced hundreds of Viennese between Saturday evening when the announcement was made and six o'clock Sunday morning when the excursion started to enroll as members of the congress. This element originally welcomed in the hope of popularizing statistics soon proved an obstacle to serious work.

2. Each meeting was unduly influenced by the public or professional opinion of the country or city where it met and from which two-thirds or more of the voting members came.

3. The congresses changed greatly in membership from meeting to meeting and so were unable to maintain continuity.

4. The relation of the congresses to the statistical offices of the various governments was quasi-official and thus extraneous considerations about the attitude of the home government were frequently in mind if not often mentioned. In other words, the congresses and the permanent commission were drawn into politics.

5. The permanent commission on the other hand kept out those economists and statists, however competent and distinguished, who did not coöperate in producing the Statistique Internationale de l'Europe.

[1] Stat. Society of London, *Jubilee Volume* (1885) p. 343.
[2] *Annales de Demog.*, 1879, pp. 303–304.

Just as this series of international gatherings came to an end a second series of meetings, which before that date had been devoted solely to hygiene, extended its field to include population statistics and vital statistics or demography as those two subjects have come to be called. This International Congress of Hygiene and Demography has now met fifteen times; the last session was held at Washington in 1912. When it adjourned it seemed likely to meet in 1917 in Russia. This international organization has probably been wrecked upon the obstacles arising from the World War as the international statistical congresses were wrecked upon the international jealousies engendered by the Franco-Prussian War.

The Congresses of Hygiene and Demography, however, failed to satisfy many European statists who had tasted the advantages of international intercourse. When the 25th anniversary of the founding of the Paris Statistical Society and the 50th anniversary of the founding of the London Statistical Society, the latter the oldest and most important of the major statistical societies, were celebrated in the same year, many thought that these gatherings might furnish the opportunity for reknitting international relations. Indeed a discussion of just this possibility was put forward by the London Statistical Society as a main part of the program for its jubilee meeting to which as well as to the earlier Paris meeting the various countries were invited to send delegates. Most countries accepted but no German delegates official or unofficial attended either gathering. The English committee wisely had asked the ablest European advocate of internationalism in statistics, Neumann-Spallart of Vienna, to sketch the history of international statistics and draft a constitution for a new international society avoiding the mistakes and profiting by the lessons of the statistical congresses and the permanent commission. In this way the International Statistical Institute was launched as a free or unofficial international society, limited to two hundred regular and twenty honorary members, not more than one-fifth of them to come from any one country or federation, allowing in its papers and debates any one of four languages, English, French, German and Italian, and not attempting in any way to commit a government by its votes or recommendations.

This organization met biennially in each odd year between 1885 and the World War and at the 1913 meeting in Vienna accepted the invitation of Belgium to assemble at Brussels in 1915. That gathering was delayed for eight years and after the armistice there was no little talk of disbanding the Institute. Fortunately its secretary was a citizen of a neutral country and had been able to keep up its scientific work and

its publications. After the Genoa Conference the need of the League of Nations for fuller and more trustworthy and comparable figures about the economic condition of Europe aroused Belgium to renew its invitation and the Institute gladly accepted it. The meeting at Brussels was highly successful and at that session the Italian Government invited the Institute to hold its sixteenth session at Rome in 1925. The invitation was gratefully accepted and thus, after an eight year interval, the series of biennial meetings has again been launched and under good auspices.

II

The United States apparently failed to coöperate officially with the London International Exhibition of 1851. No appropriation for the American section or for the expenses of American exhibitors seems to have been made. Notwithstanding this divergence from the practice of European governments, the plan for an exhibition became widely known in the United States and many Americans sent exhibits. An adequate amount of space was allotted and a commission to take charge of the interests of the American exhibitors appointed. But preparations dragged and the English press began to comment sharply upon the unreadiness of the American section. Through the liberality of an American merchant prince, Mr. George Peabody, then living in London, this difficulty was overcome. He gave $15,000 for the expenses of organization and preparation of the American section and with this timely help the American exhibitors made a respectable showing and received at least a fair share of the awards.

At the time of the London Exhibition Kennedy was superintendent of the American census and was sent to Europe to report on census conditions and to suggest greater uniformity between the different countries. In his report at the end of 1851 he says of his summer in Europe: "On each opportunity it gave me pleasure to . . . urge the propriety of mutual effort towards attainment of more uniform and useful statistical publications by different governments. . . . It affords me infinite gratification to state that the arrangement has been made for a general statistical congress to be held at Brussels (Belgium) during the ensuing fall."[1]

Even more definite is the claim doubtless inspired by Kennedy and made by Stuart, Secretary of the Interior. After saying that he had sent Kennedy abroad to examine European systems of statistics he adds: "He was also instructed to suggest to the proper officers of the respective governments the propriety of adopting a uniform system of

[1] *Seventh Census. Abstract*, p. 144.

classification for all nations who were in the habit of making enumerations of their population and productions. . . . I proposed that each nation should suggest its plan for taking its enumeration of the population, productions, etc., and upon an examination of their respective advantages a common system should be found for general adoption. . . . The suggestion was received with favor by many of the most eminent men of Europe and a proposition has been submitted for a congress of representatives from the principal nations of the world at Brussels for the purpose of carrying it into effect."[1]

The proposal was naïve and the assertion that the Brussels congress was called for the purpose of carrying it into effect was unfounded. Kennedy visited five European countries, conferred with official statists and found that his efforts were in line with current tendencies.[2] But there is no evidence that his proposals or mission modified the course of events.

At the fourth congress, held at London near the close of Buchanan's administration, the United States was represented officially by Judge Longstreet and the American Statistical Association by its president, Dr. Jarvis. Five other Americans including the United States minister to Great Britain, Mr. Dallas, are mentioned in the official report as members, but, as Mr. Dallas in his diary says he was not a member and as none of the other four took part in the proceedings, it is probable that Longstreet and Jarvis were the only American members to attend.

At the opening session and just after the inaugural address of the Prince Consort there was a curious interruption the account of which in the *London Morning Chronicle* of the following day[3] runs thus: . . . "Lord Brougham said that he hoped Mr. Dallas (American Minister) would not consider it impertinent on his part if he ventured to call his attention to the fact that one negro was in the room.

"The remark was received with very partial applause, the majority appearing to think that it was a somewhat uncalled-for observation.

"The colored gentleman so pointedly referred to, whose name we were informed was Mr. Delany, a fellow of the Royal Geographical Society, rose, and, addressing his Royal Highness, said: "I humbly thank his lordship for the notice which he has taken of me; he has always proved himself the sincere friend of the negro and I will simply assure your

[1] 33rd Cong., 1st Sess., Ex. Doc. No. 2, p. 513.

[2] My information about Kennedy's summer in Europe is derived mainly from his unpublished diary kept during that time, a copy of which was kindly given me years ago by his daughter, Mrs. A. E. K. Bidwell.

[3] The statements of the American Minister in his diary indicate that the report in the *Chronicle* was more accurate than that in the *London Times*.

Royal Highness and the company present that I am a man (cheers)."[1]

Dallas believed that this action by Brougham, a zealous abolitionist and one who frequently violated the amenities of debate,[2] was a premeditated and gratuitous insult to the United States, and Longstreet, the official American delegate, adopting the same view promptly sent notice of his withdrawal from the congress. Jarvis retained his membership and presented four papers. But he did little more than expound American conditions and thus made no significant contribution to the real work of the congress.

At the time of the fifth session, held in the midst of the Civil War, the Government of Prussia earnestly requested the United States to participate and, in view of its steady friendship for this country during that dark period, the request was granted and Samuel Ruggles was hurried away to Berlin armed with credentials and a letter from Kennedy.

No American attended the sixth session at Florence, but at the seventh Ruggles, reappearing as official delegate, reported to the Department of State that he encountered among the members a general desire for a meeting of the congress in the United States in the near future, and suggested that such a meeting might be arranged to follow the approaching one in St. Petersburg.

When the meeting occurred in 1872 the United States was more largely represented than it had been at any earlier session. There were three official delegates, the chief of the Bureau of Statistics in the Treasury Department, Edward Young, the ex-superintendent of the New York State Insurance Department, William Barnes, and the registrar of Providence, Edwin M. Snow. In addition seven unofficial delegates are listed but probably not all attended.

Young submitted two letters from the United States Centennial Commission inviting the congress to hold its ninth session at Washington in 1876 and added: "The invitation . . . has not yet an official character for the Congress of the United States . . . is not now in session, but it cannot be doubted, if you accept, that the representatives of our nation will cordially ratify the invitation."

[1] In the official "Report of the Proceedings" the account of this incident is watered down to this: "Having regard to the international character of the congress, Lord Brougham here called attention to the presence of Dr. Delaney, a member of the congress of African race" (page 12). For Dallas's account see "Diary of G. M. Dallas" pages 407–11. The same incident is referred to by Ruggles in his report on the statistical congress of 1869 as follows: "A statement curiously exhibiting the lights and the shadows in the history of internationality appeared at the time in the public journals, that a citizen of one of the Southern States of the American Union, Judge Longstreet, being in attendance at the London congress in 1860 under the authority of President Buchanan or some one of the heads of department in his cabinet, indignantly withdrew from the body on its admitting as a member an individual of the African race."

[2] "Chartered libertine" is what the Morning Chronicle called him in its editorial of the next day and Dallas repeats it in his Diary.

The congressional action required was recommended by President Grant and passed March 3, 1873. But this was too late for the purpose and the ninth and last session was held at Budapest. The United States seems never to have participated in the international work planned by Quetelet or in any of the meetings of the permanent commission, until the last (1878) to which Walker, although in Paris mainly for other official duties, was accredited as a delegate and at which he played a conspicuous part.

When the Statistical Society of London celebrated its fiftieth anniversary, most of the European countries except Germany sent official delegates. From the United States Walker attended as president of the American Statistical Association, participated in the discussions, spoke in response to the toast to the foreign visitors at the Lord Mayor's banquet and next day, when the Institute was formally organized, moved that the president of the London Statistical Society, Sir Rawson Rawson, be made president of the new international body. Walker was one of the four or five men most prominent at this conference, was the most distinguished statist America had produced, had been head of the censuses of 1870 and 1880, but at this meeting he carried no credentials from his government.

The men who had been prominent in the statistical congresses of 1853–76 doubtless kept in mind the wish to meet in the United States which at the close of that period had come so near to realization and during his association with them at Paris (1878) and London (1885) Walker doubtless learned, if he had not known before, of the official American invitation of 1873 and its failure. Another opportunity to secure an American meeting presented itself in connection with a proposal to celebrate at Washington in 1889 the centennial of the adoption of the Federal Constitution. At the first regular meeting of the Institute, held at Rome in 1887, a letter written by C. S. P. Hill, statistician of the Department of State and a member of the Institute, was presented. It was addressed to Bodio, then secretary of the Institute, but as director general of the Statistical Department of Italy. The letter was faulty in several particulars: it was not properly addressed, it referred to the approaching meeting as "the Congress of our International Statistical College," it conveyed no invitation either from the government or from any statistical organization, but merely mentioned a plan of celebrating in 1889 the centennial of the adoption of the American Constitution and the certainty of inaugurating a president in that year as reasons why the Institute might desire to meet in this country. No wonder that the president, Sir Rawson Rawson, after reading the letter to the assembly and thanking the writer com-

mented that the Institute could hardly at that time accept the invitation. A little later in the session after receiving a telegram from the French Secretary of Commerce and Industry, assuring the Institute of a cordial welcome in France, it was decided to hold the next meeting at Paris.

Before two years had passed American opinion was setting strongly in favor of a national exhibition to celebrate the four hundredth anniversary of the discovery of America. In connection with this exhibition it was the plan to hold many conferences of scientists and scholars. This probably accounts for the fact that in the 1888 annual meeting of the American Economic Association Steele, a friend of Walker, moved "That President Walker be requested in the name of the association to invite the International Statistical Institute to hold its annual meeting in 1892 in this country."[1] The constitution of the Institute provided for biennial rather than annual meetings and practice had set them in the odd years so that no meeting would come in 1892. From a subsequent letter by the president of the Institute to Walker[2] one would suppose that the American Statistical Association, of which also Walker was president, adopted a similar resolution but the secretary reports that no record of it can be found.

The invitation of the American Economic Association was not forwarded until seven months later and that of the American Statistical Association, the date of which is unknown, was forwarded at almost the same time. Both departed from the terms of the vote of the American Economic Association, but without materially improving the situation, for both asked the Institute to hold its next triennial meeting in 1892 within the United States. Naturally Sir Rawson Rawson was puzzled and as the Paris meeting of the Institute was only two weeks away he wired Walker stating that their meetings were biennial with the next in 1891 and asking whether this would affect the invitation. Whether a reply to this message was received before the Paris meeting and if so what its contents were I am unable to say. But from the records of the Institute it appears that the executive committee had received from the American Government a suggestion that it hold a special meeting at Washington in 1892 at the time of celebrating the centennial of the establishment of the government. A decision on this invitation was postponed and the Institute voted to hold its next meeting in 1891 at Vienna. When that time came the American situation had changed again. The commemoration of the discovery of America had been postponed until 1893 and for the first time since

[1] Am. Econ. Assn. Publications, Vol. II, p. 292.
[2] Kindly shown to me by the daughter of General Walker.

1873 an American attended in order to present an invitation or rather a group of invitations. Unlike the situation in 1873, however, the invitations were not conditional demanding subsequent American action to make them effective. Dr. E. R. L. Gould, an American member of the Institute, read an invitation from the Secretary of Foreign Affairs for the Chicago World's Fair and supporting letters from Walker, president of the American Statistical Association, and Wright, head of the United States Department of Labor, and leader of American official statists. The invitation was accepted and the Institute held its fourth biennial meeting at Chicago in 1893. From the Institute's point of view the main difficulty with the Chicago meeting was that, unlike each of the other fourteen meetings of that body, it was not held in response to an official invitation from the government of the country in which it met and consequently no official delegates attended coming at the expense of the home country. Even if the reasons for this difference were compelling, which I doubt, it had a disastrous effect upon the attendance of foreign members. At the thirteen European meetings the average attendance of members from outside the country which was the host of the Institute was forty-eight and the smallest attendance of such members at any European meeting except that of Paris, 1889, which like the organization meeting in London was boycotted by the Germans, was thirty-four. At the Chicago meeting only ten foreign members were present or about one-fifth of the average number in European meetings. It is probably safe to say that no successful American meeting of the Institute is likely to be held in the future unless the government of the United States coöperates to the extent of issuing invitations.

The European practice can be understood by an illustration which I draw from the arrangements for the meeting of the Institute in 1905 at London. At that time the writer and probably the other American members of the Institute received word from the American Secretary of State that the department had received from the British chargé d'affaires at Washington a note stating that the latter had been instructed in the name of the Royal Statistical Society to invite the writer as a member of the Institute to attend the London session. The chargé had asked the Department of State to communicate the invitation. His note added that the British Government desired members of the Institute in the service of the American Government to attend as official delegates. Accordingly, the Department of State wrote that it would be pleased to issue credentials in case that was agreeable to the writer and to the Secretary of Commerce and Labor.

The longest steps in the development of the American Census were

probably the innovations introduced at the Census of 1850 and the establishment of a Permanent Census Bureau in 1902. After the first of these, Kennedy, the superintendent of the Census, was sent abroad to gather information about European census practice and to suggest a greater degree of international uniformity in that field. After the second change a similar step was taken. As the newly appointed director of the permanent office could not go he commissioned the writer to go in his stead. Visits were paid to twelve European capitals and at each conferences were held with census officials. The International Congress of Hygiene and Demography was to meet at Brussels in the first week of September and the International Statistical Institute at Berlin in the latter part of the same month. Both of these meetings were attended. As a result of his trip the writer reported "that the Census Office would profit most from being put in touch with the past and present work of the International Statistical Institute, a body which is composed of the leading official statisticians of Europe and America and whose aim it is to strive for the improvement and coordination of statistical work in the different countries."

These two international statistical gatherings revealed to the writer, whose only previous contact with such meetings had been at the Chicago session of the Institute, the importance attached to them by European governments and officials and the elaborateness with which preparations both scientific and social were made. Thereafter he availed himself of several opportunities to attend European sessions of the Institute and was present at the London, Paris and Brussels meetings.

A new obstacle to the proper performance of their international obligations by American statists arose in 1913. An appropriation law of that year included the following sentence: "Hereafter the executive shall not extend or accept any invitation to participate in any international congress, conference, or like event without first having specific authority of law to do so."[1] The meeting of the Institute at Vienna occurred after this law was passed and accordingly Prof. R. T. Ely, an American member, who attended and was made chairman of one of the sections, was unable to get official credentials from the United States.

The International Statistical Institute established a Permanent Office in 1913 for publishing annual volumes and monthly bulletins prepared in coöperation with the leading statists of the several countries and summarizing international statistics in various fields. This work was to be supported by governmental subventions. Here, too, the United States has not followed the European practice. The Department of State twice has asked Congress for an annual grant of

[1] Chap. 149, 37 Statutes at Large 913, passed March 4, 1913.

$2,000 towards the expense of publishing these statistics but without success. With the depreciation of many continental currencies the situation became critical and there was danger that this promising line of work would have to be abandoned. In this emergency private assistance was obtained, a special grant of $5,000 being made by the Carnegie Corporation on application from the American members through Dr. Hadley, for the purpose of continuing the scientific work of the Permanent Office. The money thus granted is to be used for extending the scope and increasing the frequency of appearance of the office's monthly bulletin. A congressional resolution recently passed on the recommendation of the President has authorized an annual appropriation of not more than $2,500 for the scientific work of the Institute's Permanent Office and the deficiency appropriation bill which was lost in the filibuster at the close of the session carried an item of $2,000 for that object. There is reason, therefore, to hope that in the near future the United States will pay a fair share of the cost of this work.

The latest and most successful attempt of Americans to get an American session of one of these international statistical gatherings was made in connection with the International Congress of Hygiene and Demography. During its thirteenth session at Brussels in 1903 it appeared that many Europeans strongly desired to hold an early meeting in the United States and believed that if it could be arranged the attendance would probably be large. Thereupon the five American delegates at Brussels conferred with the American chargé and the latter cabled to the Department of State a request from them for authority to invite the congress to meet at Washington in 1909. To this came a reply sanctioning the invitation "subject to necessary future legislation to make the invitation effective." At the closing meeting of the Brussels congress the senior American delegate presented an invitation in this provisional form. Meantime the congress had decided to meet next in Berlin. Some five months after the Brussels meeting the Department of State asked Congress for authority to extend a formal invitation at the Berlin meeting and after a delay of three years Congress gave the authority. As the Berlin meeting meanwhile had been postponed, this long delay, unlike the shorter delay in 1873, was not fatal. At the Berlin meeting of 1907 the provisional American invitation was made a formal and official one and was gladly accepted. An organizing committee was selected and the Department of State sent a circular letter to the embassies and legations of the United States which said: "The International Congress of Hygiene and Demography ranks among the most important of international scien-

tific gatherings. It comprises distinguished representatives of many and diverse professions which are more or less connected with governmental administration, and it is therefore thought that representation at the congress by official delegates would not be without advantage to the government to which you are accredited." Each ambassador or minister was to extend to the government to which he was accredited and through it to the important state, provincial and municipal governments within the country an official invitation to participate. The congress at Washington was attended by one hundred and thirty-five official delegates representing the national, provincial and municipal governments of twenty-five countries. No doubt the widespread interest in hygiene and public health gave a congress devoted mainly to that field a public and official appeal to which a congress of statists alone could lay no claim. Yet it seems that the International Statistical Institute should have as strong a case for support and encouragement by the American government as by European governments, unless we admit that the professional character and security of tenure of the leading statistical positions in our Federal and State governments are far below what they are abroad. If that be true, then a further improvement in American standards of government statistics and an increased realization that statistics is really a profession are necessary before the United States can contribute as it should to this field of international work.

NOTES

THE HIGHER COST OF FOOD IN MASSACHUSETTS

By R. J. McFall

The cost of food appears to be more burdensome in the industrial sections of the eastern part of the United States than in other parts of the country. It is very commonly claimed that food costs are higher in Massachusetts than in the cities outside of New England. The purpose of this study is to examine this claim and, if it be substantiated, to ascertain the importance of such a factor in the local retail price of food and to point the way toward the most likely corrective for this one element in the high cost of living in the eastern industrial centers.

SUMMARY OF RESULTS

The results of this study show that there is truth in the claim that the general level of the retail price of food is higher in Massachusetts than in other parts of the country, that the chief cause of these higher prices lies somewhere in the local marketing systems and that the best probable corrective is improvement of the marketing machinery which distributes the food after it has arrived in the industrial centers of the State.

The extent of this retail price handicap for food in Massachusetts in the year 1922 is estimated at $61.31 per typical laborer's family when the comparison is made with the average of leading cities all over the country. It is estimated at $72.41 when the comparison is made with the cities outside of New England. This is 16.4 per cent of the cost of food in the outside cities and approximately 14 per cent of the retail cost of food in Massachusetts. When applied to the whole population of the State it amounts to a tax of at least $83,000,000 in a year.

The amount of this handicap which appears after the food is in the hands of the local city distributors is estimated to be $66.86 out of the expenditure of each typical family. This is 13 per cent of the family food budget in Massachusetts, or 15.1 per cent of the food budget outside of New England.

The causes most commonly assigned to this excess of price are the lack of local agriculture, the distance from the country's major sources of food production, and the charges for moving food into the markets

of the State. In so far as the system which handles food within these cities pays an extra price for what it receives from the producers and outside distributors, this extra price may be assigned to such causes as those mentioned above. In so far as an extra price for food appears after the food has reached the distributors in these cities, this extra price is not due to any of the causes mentioned above, nor to any other cause outside the cities themselves. The part of the extra cost of food which appears after the food is in the hands of the city distributors rests in the charge for local distribution. It may be assigned to the cost of such distribution, for no evidence appears that the food distributors of Massachusetts are unusually prosperous as individuals.

The difference between city wholesale prices in Massachusetts and outside of New England is estimated at only 1.29 per cent of the latter. This is a small difference. It leaves but little room for the effect of lack of local agriculture, the distance from the major producing regions and for the "tolls" of the national distributors. It discounts the importance of transportation costs. It shows that the major part of the retail price handicap of Massachusetts must be ascribed to the systems of distributing food within its own cities.

THE RETAIL PRICE HANDICAP

The study by the Bureau of Labor Statistics, on which our comparisons of retail food costs are primarily based, was conducted in coöperation with the National War Labor Board in 1918 and early 1919, and is the most recent and dependable work upon which such comparisons may be made. It covered 92 cities in 42 states, the cities varying in size from New York to country towns of a few thousand population, in order to show living conditions in representative localities all over the country.[1] The findings are actually of food expenditures, but it will be shown that they afford an excellent basis for comparing the general level of food prices in Massachusetts with the level of prices of food in the average city outside of New England.

The data for all the 92 cities were published by the Bureau, grouping the families according to size of income in each city and showing the expenditures for food and the other major items in the budget for each income group. This presentation of these data is interesting in that it shows the food costs to be higher in Massachusetts than in the majority of other cities. It is almost impossible, however, to derive from it any quantitative comparison between cities.

Shortly after the publication of these detailed data the Bureau also published summaries and conclusions from the original data made by

[1] *Monthly Review*, United States Bureau of Labor Statistics, May, 1919, p. 147.

one of their consulting experts. One of the summary tables used the data for the cities covered most thoroughly and reduced them to a common basis so that comparisons between cities could readily be made. Using all the available data for these cities, a very careful estimate [1] was made of the average expenditure for food in typical families of the same standard size and the same total annual expenditure in each city. This typical family unit which was used was one of 3.35 equivalent adult males and $1,300 expenditure a year. The data for all the families covered by the whole report were mathematically reduced to the basis of this typical family. The results are shown in the following table in which the handicap of consumers in Massachusetts in 1918 is shown to be $54.95 per typical family, or 10.36 per cent, as compared with similar consumers outside of New England.

TABLE I

COMPARISON OF RETAIL FOOD COSTS IN VARIOUS CITIES FOR FAMILIES OF THE SAME SIZE (3.35 EQUIVALENT ADULT MALES) AND WITH THE SAME TOTAL ANNUAL EXPENDITURE ($1,300) (YEAR 1918)

Massachusetts cities		Other New England cities	
Lawrence	$602.00	Portland, Me.	$564.00
Fall River	599.00	Manchester, N. H.	564.00
Boston	593.00	Providence, R. I.	547.00
Westfield	547.00	Rutland, Vt.	547.00
Average	585.25	Bridgeport, Conn.	520.00
		Average	548.40
Cities outside of New England			
New York City	$597.00	Kansas City, Kans.	$533.00
Newark, N. J.	564.00	Virginia, Minn.	532.00
Pittsburgh, Pa.	559.00	Calumet, Mich.	532.00
Philadelphia, Pa.	555.00	Cincinnati, Ohio	529.00
Camden, N. J.	555.00	Chicago, Ill.	528.00
Scranton, Pa.	550.00	St. Louis, Mo., and East St. Louis, Ill.	524.00
Syracuse, N. Y.	534.00	Cleveland, Ohio	519.00
Buffalo, N. Y.	508.00	Columbus, Ohio	516.00
Johnston, N. Y.	547.00	Detroit, Mich.	512.00
Baltimore, Md.	550.00	Brazil, Ind.	511.00
Atlanta, Ga.	532.00	Danville, Ill.	511.00
Charlotte, N. C.	520.00	Indianapolis, Ind.	510.00
Winston-Salem, N. C.	520.00	Pana, Ill.	511.00
Richmond, Va.	545.00	Minneapolis	481.00
Charleston, S. C.	534.00	St. Paul	481.00
Huntsville, Ala.	564.00	Cripple Creek, Colo.	530.00
Meridian, Miss.	564.00	Trinidad, Colo.	530.00
New Orleans	560.00	Denver, Colo.	528.00
Houston, Texas	553.00	Seattle, Wash.	528.00
Dallas, Texas	553.00	Portland, Ore.	521.00
Birmingham, Ala.	528.00	Everett, Wash.	511.00
Memphis, Tenn.	486.00	Astoria, Ore.	511.00
Kansas City, Mo.	533.00	Los Angeles, Calif.	494.00

Average outside New England	$530.30
Massachusetts average	585.25
Average for rest of country	532.08
Massachusetts price as percentage of cities outside of New England	110.36%
Massachusetts handicap compared with cities outside of New England	$54.95

[1] *Monthly Review, ibid.*, August, 1919, p. 5.

These cities are scattered all over the country, and small as well as large are included. No separate weighting is given to the prices for each city to correspond to its relative population, for the reason that each small city represents many similar small cities while the large cities practically represent only themselves. The selection of cities is a weighting to represent the whole urban life of the country. Only four cities represent Massachusetts, but it would be difficult to select four more representatives cities for the whole urban life of the State.

These data are the best existing basis for determining the comparative retail price of food in urban Massachusetts and the urban sections of other parts of the country. The greatest possible care has been taken to eliminate the chance of factors other than the comparative costs of food in the various cities influencing the comparative expenditure for food. The use in all places of the typical family with the common low expenditure of $1,300 is the greatest assurance of an accurate comparison. Thirteen hundred dollars a year expenditure in 1918 for a family consisting of two adults and three children (3.35 equivalent adult males) left very little opportunity for a possible luxury selection of foods in any place. There is room for a higher class selection of food by the *average* family in Massachusetts than elsewhere, for the average income per person gainfully employed in the State has been estimated to have been higher than the average for the whole country. The family in Massachusetts having only $1,300 to spend was, however, below the average income here, which has been estimated at $1,755 in 1918, and was in no better position to buy food than was a similar family elsewhere.

Other data presented by the Bureau also indicate that the difference in expenditure was due practically entirely to the difference in cost of the food. Representative dietaries from eleven localities scattered throughout the country show that the consumption of meats in Boston was almost identical with the average for the whole eleven cities.[3] This eliminates the vexed question of the comparative expense of meat in the dietary. There were some slight differences in the regional selections of other food, but none presenting difficulty in comparing the cost of the food purchased in Boston with the average for the different cities in this selection. In the tables published by the Bureau exhibits are given for the expenditures of the negro population. These have been eliminated in our study.

The more detailed tables of the Bureau[2] show the expenditures for the other major items in the cost of living as well as food. Rent and

[1] *Monthly Review, ibid.*, August, 1919, p. 22.
[2] *Ibid.*, Volumes 8 and 9.

fuel, clothing and furniture are accounted for as well as miscellaneous expenditures and surplus. If, in this presentation, it was apparent that there was a much higher rent charge in one group of cities than another and the food expenditures showed a reverse comparison, the conclusion might be in doubt. The higher rentals of one group might cause a real deficiency of food. It is not possible from the published details to compute an exact average expenditure in these other items. However, after comparing these details for the $1,200–$1,500 income group for Massachusetts cities and those outside New England, it is evident that variations in the comparatively inexorable dwelling costs of rent and fuel do not account for the difference in food costs. This may or may not apply to other geographical groupings of cities, but rent and fuel were not particularly low in the cities of Massachusetts. The difference in food costs appears chiefly in the more elastic and less vitally necessary items of miscellaneous expenditures and surplus which were both low in Massachusetts.

Fortunately, data are available[1] for eleven cities scattered throughout the country showing the expenditure in each city for an approximately equal number of calories—a number of calories just sufficient to provide an adequate dietary. In selecting these data, no restricted income group is considered. Only white families were considered, but otherwise the selection was made solely on the basis of the amount of food purchased. Thus each city is represented by white families using similar amounts of food in terms of energy units. On this basis the cost of food in Boston was 10.65 per cent higher than the average cost in the following ten cities: New York, Chicago, San Francisco, Seattle, Minneapolis, Denver, St. Louis, New Orleans, Atlanta and St. Paul.

In this comparison variations in rent and other demands upon the family purse have no effect upon the conclusion. Such variations in expenditures are balanced by variations in income. The amount of food in terms of the only available common denominator is constant. The variation in expenditure for food must almost certainly be due entirely to difference in price.

It is interesting to note that this finding of a 10.65 per cent handicap is so close to the results obtained in Table I, which show a comparison of 10.36 per cent. It would appear not only that the food expenditures in Table I represent fairly the comparative levels of retail food prices, but that the margin of error in the comparison is small.

These data represent the state of affairs in 1918. They cover a twelve months' period substantially that of the calendar year 1918. The regular index numbers of the Bureau show the change in the retail

[1] *Monthly Review, ibid.*, August, 1919.

price of food in all the leading cities for each year and each month since the beginning of 1913. In these Massachusetts is represented by Boston and Fall River. The outside is represented by 35 cities.[1]

It is evident from these index numbers that the year 1918 was in a period during which Massachusetts experienced a relatively smaller handicap than at any other time in the past decade. The war boom in many other parts of the country had caused prices to advance more rapidly elsewhere than here until 1918. Since the war the tendency has been in the other direction, and on both the up and down grade of prices in this latter period the consumers of the State have been comparative losers. On the up grade in 1920 the prices in Massachusetts rose faster, and on the down grade since that year their fall has been less than the average outside of New England. As a result, the typical family food cost of the State, which was $585.25 for 1918, has fallen by only 12.12 per cent to $514.31, whereas the average for the whole country has fallen by 15.5 per cent, from $535.94 to $453, and for the cities outside of New England by 16.7 per cent, from $530.30 to $441.90. The handicap in 1922 was $61.31 per family, or 13.5 per cent, as based upon the average city in the country, and $72.41, or 16.4 per cent, as based upon the average city outside of New England. This latter handicap is 14 per cent of the food cost per family in Massachusetts. Incomplete data for 1923 indicate that this situation is continuing to grow worse.

The conclusion that a substantial handicap exists in the retail cost of food is supported by various other pieces of evidence. No one of these is as reliable for a quantitative measure as the one which we have used.

The average retail price quotations for the leading items of food on which the index numbers of the Bureau of Labor Statistics are based are frequently used for price comparisons between cities. Whatever evidence these give supports the conclusion that food prices are higher in Massachusetts than in the cities outside of New England. These price data for March 15, 1923, indicate that on that date retail food prices in Boston were 13.4 per cent higher than the average for retail prices in all cities, those of Massachusetts and other New England states being included. Were these New England cities excluded in the base, this percentage would approximate the 16.4 per cent handicap shown by the data already outlined.

These data are not gathered on a basis calculated to yield the greatest accuracy in comparing the difference in the price level of the various

[1] Mimeographed compilation of the Bureau of Labor Statistics, "Index numbers of retail prices of food in specified cities, 1913 to May, 1923, by months and years."

cities. However, since they give substantially the same results as the data which we have been considering while being worked out on such totally different systems, they may be considered to be highly corroborative evidence of the accuracy of the findings based upon the other data.

All the evidence presented and that available from various sources support the conclusion that a retail price handicap exists in the purchase of food in Massachusetts cities. The best evidence available establishes the amount of this handicap at somewhat over seventy dollars per typical family of very modest means in 1922, an amount which is continuing to increase. This is 16.4 per cent of the cost of food in cities outside of New England, and is 14 per cent of the cost of food in Massachusetts.

THE DIFFERENCE IN WHOLESALE PRICES

All available evidence indicates that the difference in wholesale prices in Massachusetts cities and in cities outside of New England is very small. After weighting each class of food to allow for its relative importance in the budget, we have estimated that food costs at wholesale only 1.29 per cent more in Massachusetts than in the other cities.

This fact may seem surprising until we consider that freight costs are small as compared with the final wholesale price of food in our cities. Freight on meat shipped to Boston from Chicago is only 87 cents per cwt. On flour it is only 45.5 cents. On meat the freight rate from Chicago to Massachusetts points is exactly the same as to New York, and only one-fiftieth of a cent a pound more than to Philadelphia.[1]

As a rule, wholesale prices are very much more flexible than retail prices. It is very difficult for wholesale prices in any one place to vary much from those in other places by an amount much greater than the transportation costs. They also change very much more quickly than retail prices in a change of economic conditions in point of time. This characteristic of wholesale prices keeps the wholesale level closely in line with outside markets, while something is lacking to keep the retail prices in close correspondence with the outside.

The fact that all available data show that there is very little regional variation in wholesale prices makes it less necessary to secure data from a wide selection of cities in order to present a true picture of the wholesale price comparison. The greater the variation there is in the sizes of things to be averaged, the greater number must be selected to insure a true average. All possible care has been taken to get whole-

[1] Railway rate data supplied by Bureau of Agricultural Economics, U. S. D. A.

sale price data from representative cities and to ascertain the degree to which these cities are representative.

All evidence which exists points to the conclusion that the wholesale prices paid in the larger cities are on the whole lower than those in the adjacent smaller cities. This is in direct contrast to the retail price situation, where prices tend to be higher in the larger centers.

It is impossible to give exact quantitative evidence on this point, but various matters lead to this general conclusion. A very large part of the food used all over the country goes through the major channels of commerce rather than by a direct route from farm to kitchen. Two-thirds of our meat is slaughtered in federally inspected establishments.[1] This is the meat which goes into interstate commerce and is used in the cities rather than in the non-urban sections. Some exceptions exist to this rule, especially in New England. Since less than two-thirds of the national population is urban, it is evident that this "packer" meat is predominant in the cities and that "local" meat has very little influence on city prices throughout the country. Eighty per cent of all our food is estimated to be manufactured or processed to some extent.[2] The greater part of such foods as they leave the factory is more available to the large than to the small city, and competition in their first sale is keener in the large cities. Freight rates are practically never higher to the larger centers than to neighboring small places and the latter frequently pay an additional charge.

In the case of meat, which is one-third of the food budget, we have trade information in our files which establishes the fact that wholesale prices are actually lower in the large cities. Sugar prices vary only with the transportation charges, which tend to be higher in the small cities. Flour and dried and canned fruits and vegetables are practically on the same basis. Fresh fruits and vegetables during a comparatively short season of local production may be cheaper in the smaller places, but throughout the rest of the year such supplies are largely shipped in on the same basis as to large cities, or are even reshipped at added expense. The Joint Commission of Agricultural Inquiry (67th Congress) said regarding fruits and vegetables: "People in cities now draw their major supplies of these commodities from distant points. . . . There has been a very noticeable tendency toward the concentration of fruit and vegetable acreage in certain specialized restricted districts."[3] The situation regarding milk is different, this food being highly perishable and on a comparatively

[1] Annual address of the President of the Institute of American Meat Packers, 1923, p. 18.
[2] Macklin, Theodore. *Efficient Marketing in Agriculture*, p. 269.
[3] *Joint Commission of Agricultural Inquiry* (67th Congress), Part IV, p. 68.

local trading basis the year round. For this commodity we have obtained a list of cities as a basis for wholesale prices which is similar to that used for retail costs.

We believe that any error which arises from the selection of cities used as a basis for wholesale comparisons is very small, but that such as it is it tends to a more conservative final result.

MEATS

The most important single class of foodstuffs from the standpoint of its share in the food budget is that of meats, including a small amount of poultry and fish. The proportion of the total average food expenditure devoted to this class of food is from 33–40 per cent according to different estimates. It has been impossible to obtain exactly comparable wholesale quotations for meats from a wide range of cities. Trade data, which cannot be incorporated into an exact quantitative comparison, show prices in Chicago to be somewhat lower than the average for such cities as Washington, Cleveland, Indianapolis, Detroit, Des Moines, Fort Worth, San Antonio, Atlanta, and San Francisco. It is not unfair to use Boston as the sole representative for Massachusetts, since available trade data establish the fact that differences between the wholesale prices in Boston and those in other Massachusetts cities are comparable to those between the large and small cities outside of New England.

Data furnished us by the United States Department of Agriculture covering wholesale prices for different classes of meats in 1922, in Boston, New York, Philadelphia and Chicago, show the average price of meats in Boston to be only 98.57 per cent of the average in these outside cities. In the light of the foregoing remarks this may be taken as a conservative presentation of the relation between meat prices in Massachusetts cities and prices for similar meats in cities outside of New England.

We have used the quotations for such sub-classes of beef, veal, pork, mutton and lamb, as were quoted consistently for all cities, and the results show what average prices had to be paid at wholesale in each of these cities for the same kinds of meats. The comparison was computed on both a simple and a weighted basis, and the difference in result for these two methods was only .04 per cent. The weighting used was the same for Boston and for other cities. These weightings were based upon the average relative use of each class of meat according to the best information available. An experimental computation was made using as a variety weight for Boston the proportion of each variety consumed in the North Atlantic states, and for the other cities

that consumed in the country as whole. This result was not distinguishable from that of the method using one standard weighting.

The result, which shows Boston's wholesale meat price to be approximately one and a half per cent below the average, is surprising. The previous reference to the small part which freight rates play prepares us for only a small difference, but does not account for lower prices in Massachusetts. We have gone through the matter carefully with leaders in the trade and find this result fully supported, both as to fact and a probable cause. The probable cause is a very keen competition between packers, or national distributors, large and small in Boston and its vicinity. When other markets in the country are oversupplied, products are reshipped to save a glut. New England markets lack such an outlet, not having the advantage of other markets more distant from the original sources of supplies. On the other hand, they are subjected to considerable such "dumping" from the other markets.

DAIRY PRODUCTS

The foodstuffs of next relative importance in the budget are dairy products. These we shall divide into fresh milk and other dairy products.

FRESH MILK

The United States Department of Agriculture publishes data currently on the price of milk in various cities reducing the welter of differences in quoting prices to that for milk of 3.5 per cent butter fat f. o. b. each city. We have selected a list of cities comparable to that used in Table I.

Boston, Fall River, Lawrence and Springfield are taken as the representatives of Massachusetts cities, and thirty-five cities of comparable sizes and location to those used in our retail comparison are taken to represent the urban cities outside of New England. Monthly quotations for December, 1922, to July, 1923, inclusive, are used. The results, after making careful corrections for the different ways of selling milk, show that the wholesale price of milk in Massachusetts for that period was 15.5 per cent above the price level outside of New England.

OTHER DAIRY PRODUCTS

Butter is the most important product in this class. Our basis for comparison in this case is monthly average prices for 92 score butter during the year 1922 at Boston, New York, Chicago and Philadelphia, as supplied by the United States Department of Agriculture. The average price for Boston is 40.66 cents per pound, and for the other

cities 40.33 cents per pound. Boston's price is, accordingly, 100.8 per cent of the average price of other cities. Since cheese and other dairy products occupy such a relatively small part of the food budget, butter prices are taken to represent this class. Boston's price is taken to represent Massachusetts, since prices all over New England are made on this basis.

EGGS

Our data for eggs are comparable to those for butter, excepting that the average for the year 1921 is taken as the basis. Boston's average price was 41 cents and the average for New York, Philadelphia and Chicago was 40 cents. The Massachusetts price was, accordingly, 2.5 per cent above that of the outside cities.

CEREALS

The most important element of this class is wheat flour. Price data for this commodity, accordingly, were compiled from the weekly quotations given in the *Northwestern Miller* in the early part of the year 1923. The cities covered in its price table are: Chicago, Minneapolis, Kansas City, St. Louis, New York, Baltimore, Philadelphia, Columbus, Nashville, Seattle, San Francisco and Boston. Using several grades of flour, prices in Boston were found to be 7.7 per cent above the average of the cities outside of New England. Boston's price for flour is the Massachusetts price for flour. The freight rate from the chief milling centers to all Massachusetts cities is the same. The larger centers are supplied by direct shipment on a common price system, car-lot prices being the same at all points in the State. Flour at 7.7 per cent above outside prices is, accordingly, taken to represent the comparison between Massachusetts wholesale prices for cereals and the average for the cities outside of New England.

SUGAR

Sugar prices throughout the country are on a very systematic basis. This is inevitable, since sugar is a completely standardized article. Its price system is based upon the prices at the cane sugar refineries on the Atlantic Coast. These prices are substantially the same. The price at any other than a refinery point is the price at the nearest refinery plus the cost of transportation. Beet sugar sells at a certain standard differential below cane.[1]

In such conditions recourse must simply be had to comparing the price at Boston with the basic refinery price plus freight to some point

[1] Letter to R. J. McFall from Willett and Gray.
Letter to R. J. McFall from U. S. Sugar Association, July 9, 1923.

approximately in the center of the population outside of New England. In using such a method of comparison the influence of prices in the smaller Massachusetts cities, which are slightly higher than those in Boston, will offset the tendency toward greater freight charges in the smaller cities in the interior of the country. The average price of sugar in 1922 at the refineries of the East was 5.9 cents per pound which, as we have seen, was the same for all the seaboard refinery centers including Boston. The average price for cities outside of New England we take to be their price plus freight from New York to Chicago. This freight being 56½ cents per cwt. to Chicago, which is approximately the center of population outside of New England, the price in Massachusetts is seen to be 8.74 per cent below the outside level.

FRUITS

Fruits are a class of food on which it is difficult to get average prices for comparable classes and grades at different places. Citrus fruits are the most highly standardized of any, both in quality of product and method of sale. Average prices for all oranges sold by the California Fruit Growers' Exchange at the leading markets in 1922 show Boston's average price per box to be $6.05, and the average for thirty-two United States cities including New England cities to be $5.70. This shows Boston's average price for this fruit to be 104.49 per cent of the average of the whole country.

We also have prices of apples[1] in Boston, New York, Baltimore, Chicago and St. Louis. These are for definite varieties and largely for a specific grade. Boxed apples are given separately from barreled apples. The price of the latter in Boston was 103.01 per cent of the average for all cities. In the case of boxed apples, the Boston price was lower than the average, appearing as 98.37 per cent.

The percentages for these fruits are reduced to a common basis by using the carload arrivals of each as reported by the Boston Chamber of Commerce as weightings in computing the average. The result shows the wholesale price of fruit in Boston to be 103.1 per cent of the average of all the cities. Considering the fact that New England is included in the average of all the cities, it is necessary to reduce this figure slightly to make it show a comparison with cities strictly outside of New England. 102.75 per cent is taken as the final comparative figure.

VEGETABLES

In the case of vegetables we were able to obtain comparable prices for potatoes, onions, tomatoes and lettuce.[1] While this is by no

[1] Data supplied by Bureau of Agricultural Economics, U. S. D. A.

means a complete list of vegetables, it is highly representative and includes the most important types. The same cities are covered as in the case of fruit. These commodities are not very well standardized, but the prices appear to be worthy of comparison, since in type of package, variety and other respects, these vegetables were sold as uniformly in the various cities as such vegetables apparently could be sold.

In the case of potatoes, both early and late varieties are included, and every month of the year 1922 but June is represented. Boston's price appears as 97.85 per cent of the average of all the recorded cities. Onions sold at a higher price in Boston than the average city in spite of the nearness of the Connecticut Valley onion fields. Eastern and middle western yellow varieties take the largest share in the price comparison, but Texas yellow Bermudas appear for three months. Boston's price for onions was 106.7 per cent of the average of all recorded cities. Tomatoes appear as 91.25 per cent of the average of recorded cities. The prices cover only tomatoes shipped in interstate trade. Lettuce also appears lower in Boston. Here again the product from interstate trade is covered. Boston's price is 98.76 per cent of the average of all cities.

These various vegetable prices are brought together by using a similar system of weighting to that used in the case of fruit. Since the result is so close to 100 per cent, no adjustment is necessary to offset Boston in the base average. The average price in Massachusetts is found to 98.7 per cent of that in outside cities.

COMBINED WHOLESALE PRICE COMPARISON

It now remains to assemble these percentages for the different major classes of foods into one general comparison. To do this due care must be taken to allow for the different importance of each class of food in the normal consumer's budget. On this account the proper weighting must be given to each food class. The weighting which we have used is one computed by Professor H. C. Sherman[1], from the data of the Bureau of Labor Statistics—whose data we have used for our retail cost comparison. In this the class of meats, including poultry and fish, is found to account for 33.8 per cent of the food budget; eggs, 5.14 per cent; milk, 6.52 per cent; other dairy products, 12.46 per cent; cereal products, 9.57 per cent; sugar, 5.34 per cent; fruits, 5.05 per cent; vegetables, 9.72 per cent. The miscellaneous items of the food budget, such as tea, coffee and condiments, cannot be computed. Since many of them are imported, the position of Massachusetts on

[1] Sherman, H. C. *Chemistry of Food and Nutrition* (revised edition), pp. 386–392.

the seaboard gives a tendency for their prices to follow that of sugar rather than that of flour. It being impossible to compute any quantitative comparison for such minor items, we draw our conclusions from the major known items. The following table shows the method of working out the final wholesale comparison and shows that the wholesale prices of food in Massachusetts cities are only 1.29 per cent above the average level for cities outside of New England.

TABLE II

WHOLESALE PRICE COMPARISON—WEIGHTED SUMMARY

Class of food	Ratio of wholesale price in Massachusetts cities to average of other cities	Percentage weight in food budget	Weighted relative
Meats and fish	98.57	33.80	33.317
Eggs	102.50	5.14	5.218
Milk	115.30	6.52	7.531
Butter and cheese	100.83	12.46	12.563
Cereals	107.76	9.57	10.313
Sugar and molasses	91.26	5.34	4.873
Fruits	103.50	5.05	5.227
Vegetables	99.70	9.72	9.691
Total	101.29	87.60	88.735

The amount of the State's handicap in the retail cost of food was estimated at 16.4 per cent in 1922. The comparison of wholesale prices, using data as near as possible to the 1922 average, shows a handicap of only 1.29 per cent. This latter shows as nearly as possible the price situation when the food enters the systems of local city distribution. The retail handicap shows the price situation when food leaves the systems of city distribution. Wholesale prices and retail prices are, of course, at different levels. However, the accepted business principle is that any increase in wholesale prices must in general be carried over in full proportion, rather than amount, into retail prices. Relying on this principle, we subtract the two percentages and reduce the difference to its percentage of the retail figure to find the proportion of the retail price handicap not due to wholesale prices which reflect factors outside of the local cities. The difference between 116.4 per cent and 101.29 per cent is 15.1 per cent; 15.1 per cent is 13.0 per cent of 116.4 per cent. Massachusetts cities then have a handicap in the cost of retail food due to nothing outside of these cities nor to their relations to outside sources or systems of supply which amounts to $66.86, or 15.1 per cent, as based upon the cities outside of New England, or 13.0 per cent of their own food budget. This price difference originates in the local systems of city food distribution.

CONCLUSION

The object of this study was primarily to shed light upon the food problem of the State of Massachusetts. Massachusetts is an industrial state containing many cities. The tendency appears to be toward relatively high retail prices for food in all large cities—retail prices not accounted for by higher wholesale prices. The food problems of Massachusetts appear to a certain extent to be typical of all industrial regions with congested populations. However, the problem seems to be more acute in Massachusetts than in many others of these regions. The outside cities used for purposes of comparison include New York and all the other large cities. To the extent that these other large cities are also burdened with a food price handicap over smaller cities due to their own systems of distribution, the price handicap of Massachusetts is underestimated as compared with less crowded districts. Had our comparisons been made with the smaller cities alone outside of New England, the handicap would have appeared as much greater than what we have shown.

In so far as this situation has a general tendency for all large industrial centers, the issue seems to be either a scattering of the industrial life of the nation into smaller units or a corrective for the high costs of urban food distribution. The degree to which such issue will be pushed depends, of course, upon the presence, or absence, of other factors playing upon the situation.

For Massachusetts the situation has an especially serious bearing. Its food handicap is comparatively very acute. Its industrial position is due partially to the chance factors of early settlement rather than the presence of power or raw materials, which are largely brought into the region by rail. A community in such a situation must dispense with all unnecessary handicaps. The presumption is that this handicap is unnecessary, since the comparison has been with cities of comparable size in the other sections of the country. As a working hypothesis it may be assumed that food may be distributed as cheaply in Massachusetts as in other urban sections.

We do not advocate government regulation. Such a policy might produce more harm than good. We do advocate increasing study of the concrete problems involved with a view to the possible introduction of economies in the merchandising of food. The distribution of food is now being studied in the interest of the farmers. It should be studied still more intensively in the interest of the consumers. The problem bears upon the consumers in all sections of the country, but it is vitally acute in the State of Massachusetts. Fortunately for the

State, it does not have to go beyond its own borders to find the cause of its high food prices. The careful study of the local systems of food distribution within its own cities presents the possibility for a solution of a very material part of the cost of living in Massachusetts.

WHOLESALE PRICES, 1801–1840

By Alvin H. Hansen, *University of Minnesota*

In the October, 1923, *Supplement to the Review of Economic Statistics*, Mr. N. J. Silberling presents a new index of British wholesale prices for the period 1779–1850. In this article he gives a chart in which he compares his index with the index of wholesale prices for the United States constructed by Harold V. Roelse for the period 1791–1801 and by the present writer for the period 1801–40. Mr. Silberling's index is the simple geometric mean, while Mr. Roelse's and my own were calculated by the simple arithmetic method. I have since had my index recalculated by the geometric mean and the median methods. Table I gives the relatives for the three methods.

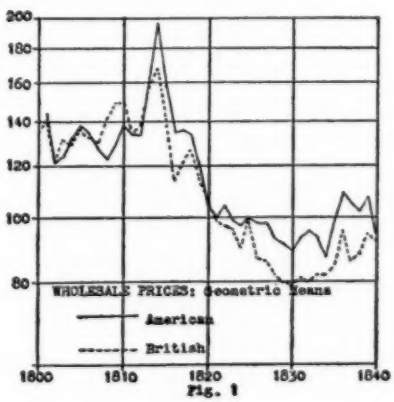


Fig. 1

In Figure 1 the curves for British and American wholesale prices are plotted together. The geometric mean is used in both cases. British prices did not rise as high as American prices in 1814, but fell as rapidly in 1815–16. In 1817–18 the rise was more pronounced than in the United States. Following the crisis of 1825 British prices fell lower than American prices. On the other hand, the price upheaval of the thirties was more pronounced in this country. The close sim-

TABLE I

WHOLESALE PRICES IN THE UNITED STATES, 1801–1840
(Base year: 1825)

Year	Geometric mean	Median	Arithmetic mean
1801	144	146	156
1802	121	120	128
1803	124	128	131
1804	133	139	141
1805	138	136	145
1806	135	136	142
1807	128	127	134
1808	123	123	131
1809	130	132	137
1810	138	136	150
1811	134	135	146
1812	134	130	148
1813	161	150	172
1814	196	187	215
1815	158	161	169
1816	135	132	144
1817	136	138	145
1818	134	137	142
1819	122	123	125
1820	104	106	107
1821	100	100	102
1822	104	103	105
1823	99	100	100
1824	97	96	99
1825	100	100	100
1826	98	98	99
1827	98	99	100
1828	93	97	95
1829	91	93	94
1830	89	93	91
1831	93	97	97
1832	95	98	98
1833	93	99	97
1834	87	92	90
1835	99	101	104
1836	109	115	116
1837	105	110	115
1838	102	105	110
1839	107	116	115
1840	92	92	98

ilarity in the two curves, not only as to trends, but particularly with reference to the short-run cycles, is indeed striking.

Of the total number of commodities used in my index,[1] the various groups of commodities, corresponding roughly to the classification used by the Bureau of Labor Statistics, constituted the following percentages:

TABLE II

	Period 1801–1825	Period 1825–1840
Foods	35.4	49.2
Farm products	13.9	20.6
Clothes and clothing	3.8	4.8
Fuel and light	2.5	3.2
Metals and metal products	8.9	4.8
Building materials	16.5	7.9
Drugs and chemicals	7.6	4.8
Miscellaneous	11.4	4.8
Total	100.0	100.1

[1] *Cf.* earlier article by the present writer on "Wholesale Prices for the United States, 1801–1840," *Quarterly Publications of the American Statistical Association*, December, 1915, pp. 804–812.

Separate indexes of foods and farm products have also been calculated by the geometric mean method. In connection with another study I have constructed weighted index numbers of both foods and farm products for this period. While these weighted index numbers are not for identically the same commodities as the unweighted index numbers, nevertheless, I have included them in the table given below for purposes of comparison.

Figure 2 gives the geometric means for *all commodities, foods* and *farm products.* Food prices rose considerably above the general prices in 1805–6. Foods did not rise as high as the all-commodity index in

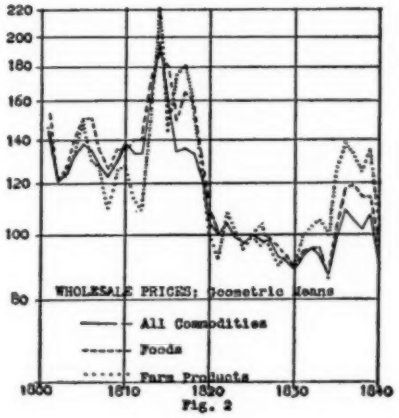


Fig. 2

1814, nor was the drop as precipitous in the following year. The food index rose to a point considerably above the general index in 1817–18, and again in 1836–9.

The farm products index went above general prices in 1805, dropped to low points relative to the general price level in 1808 and 1812, rose sharply to points above the general price level in 1814 and 1817, and suffered a tremendous collapse from 1818 to 1821. During the wildcat banking upheaval of the thirties, farm product prices rose considerably above the general price level.

During the period of the late thirties the weighted index numbers for both farm products and foods given in Tables III and IV reached exceptionally high levels. This was due to the fact that the heavily weighted commodities experienced extraordinary price increases in this period. Wheat rose from the index number 115 in 1834 to 209 in 1838, while corn rose from 120 in 1834 to 190 in 1837. The relatives for flour increased from 97 in 1834 to 178 in 1837. Pork prices increased from 96 in 1834 to 169 in 1836, while the index number for beef increased from 101 in 1834 to 191 in 1838.

TABLE III

WHOLESALE PRICES OF FOODS

(Base year: 1825)

Year	Unweighted geometric mean { 18 commodities, 1801–1825* } { 20 commodities, 1825–1840† }	Weighted arithmetic mean (12 commodities‡)
1801	153	157
1802	121	119
1803	127	124
1804	140	137
1805	150	151
1806	150	144
1807	135	131
1808	127	111
1809	136	122
1810	135	127
1811	140	130
1812	140	130
1813	169	164
1814	189	166
1815	180	156
1816	149	143
1817	165	166
1818	160	159
1819	133	134
1820	110	104
1821	100	94
1822	104	107
1823	99	104
1824	97	95
1825	100	100
1826	96	92
1827	99	101
1828	96	101
1829	93	106
1830	90	94
1831	94	97
1832	95	99
1833	95	98
1834	86	91
1835	102	110
1836	118	124
1837	119	131
1838	115	140
1839	115	136
1840	97	113

* The eighteen commodities are as follows: flour, corn meal, bread, mackerel, pork, beef, bacon, lard, butter, cheese, raisins, lemons, coffee, tea, molasses, sugar, salt, wine. Several unimportant foods included in the general index are here omitted.

† The twenty commodities are: wheat flour, rye flour, corn meal, cod, mackerel, pork mess, pork prime, beef mess, beef prime, hams, lard, butter, cheese, raisins, coffee, tea, molasses, sugar, salt, wine. Unimportant series are omitted

‡ The twelve commodities used in the weighted index are: beef, pork, butter, lard, flour, fish (cod and mackerel), cheese, fruit (currents, prunes, figs, raisins), molasses, sugar, coffee, tea. The weights used are roughly according to the expenditures of 232 families as given on page 62 of Senate Report No. 1394, Part I.

TABLE IV

WHOLESALE PRICES OF FARM PRODUCTS

(Base year: 1825)

Year	Unweighted geometric mean { 8 commodities, 1801–1825* } { 9 commodities, 1825–1840† }	Weighted arithmetic mean (9 commodities‡)
1801	146	162
1802	123	122
1803	123	123
1804	133	137
1805	149	162
1806	131	142
1807	128	139
1808	110	115
1809	126	124
1810	128	140
1811	114	134
1812	109	130
1813	156	176
1814	221	211
1815	144	160
1816	174	182
1817	180	211
1818	156	172
1819	130	134
1820	101	101
1821	92	88
1822	108	107
1823	101	98
1824	95	90
1825	100	100
1826	104	103
1827	96	95
1828	90	94
1829	93	99
1830	90	93
1831	99	103
1832	103	105
1833	105	111
1834	100	103
1835	125	135
1836	138	151
1837	134	152
1838	125	146
1839	135	143
1840	101	105

* The eight commodities for the period 1801–25 are: cotton, barley, rye, oats, corn, rice, hides, tobacco. A few minor agricultural commodities included in the general index are omitted.

† The nine commodities for the period 1825–40 are: cotton, wheat, rye, oats, corn, rice, hides, wool, tobacco.

‡ The nine commodities are: cotton, corn, wheat (or wheat flour), oats, barley, rye, pork, beef, tobacco. The weights used are roughly the value product in 1859 as given on page 108 of Senate Report No. 1394, Part I. No really satisfactory weights are available.

MONTHLY PIG IRON PRODUCTION, 1884–1902: AN ANALYSIS OF A PROBLEM IN CORRELATION

By Holbrook Working

For the purpose of obtaining monthly data on pig iron production prior to 1903, Margaret G. Myers recently made an interesting use of the Pearsonian method of correlation.[1] A further analysis of the results obtained by Miss Myers makes possible an improvement in the

[1] *Cf.* this Journal, June, 1922, pp. 247–249.

production figures for the period prior to 1902, and also serves to illustrate some points of importance regarding the use of the Pearsonian method of correlation. Miss Myers' study offers an illustration of the use of a regression equation obtainable from a correlation study, while the present paper gives a further interpretation of this equation.

The importance of pig iron production as a measure of business activity is generally recognized. The monthly figures on pig iron production have been available since 1902 from the tabulations prepared by the *Iron Age*. More recently these figures, usually beginning with 1903, have been made more conveniently available to economists in the publications of the Harvard Committee on Economic Research and elsewhere. Prior to 1902 the only figures available have been the reports of the weekly capacity of furnaces in blast at the first of each month, as compiled by the *Iron Age*.

As a basis for converting these weekly capacity figures into production per month, Miss Myers took data for the period from 1903 to 1909, during which time both sets of figures were available, and made a correlation of weekly capacity of furnaces in blast at the first of each month with the production for the month. The coefficient of correla-

TABLE I

WEEKLY CAPACITY OF FURNACES IN BLAST AT THE FIRST OF EACH MONTH, 1884–1902*

Month	1884	1885	1886	1887	1888	1889	1890	1891	1892	1893
January	74	58	84	114	121	142	162	155	177	165
February	....	63	85	121	104	138	163	135	176	162
March	....	70	85	123	97	139	168	124	182	168
April	....	67	98	125	101	138	168	104	175	170
May	77	67	106	127	107	134	169	106	168	173
June	73	65	109	89	108	126	168	137	162	166
July	75	64	108	98	102	131	163	160	159	146
August	70	65	106	100	108	131	154	159	147	102
September	66	63	102	121	115	133	159	159	142	78
October	64	64	107	129	119	139	166	168	149	68
November	66	68	109	130	128	153	165	176	162	75
December	67	77	111	128	137	156	171	178	167	94

Month	1894	1895	1896	1897	1898	1899	1900	1901	1902
January	95	164	202	154	221	237	287	243	292
February	96	158	192	158	224	233	290	278	325
March	106	153	184	165	229	224	286	285	316
April	123	154	182	168	228	241	272	289	331
May	106	152	185	167	228	245	287	294	338
June	59	153	177	165	219	249	289	307	337
July	82	167	174	161	211	257	275	304	343
August	112	176	151	161	200	261	236	297	329
September	145	190	123	181	207	262	224	293	328
October	146	197	106	195	210	272	215	301	338
November	158	212	119	209	223	281	207	313	330
December	164	212	137	221	230	289	222	317	337

*Data taken from "Iron Age."

tion was +.975 and the equation of regression of monthly production on weekly capacity at the first of the month was $y=8.1+4.34x$, y representing production for the month and x weekly capacity of furnaces in blast at the first of the month, the unit being thousands of tons. Being interpreted, this equation means that on the average the production during a month was 4.34 times the weekly capacity of furnaces in blast at the first of the month, plus 8,100 tons in addition.

The conclusion that the production during each month averaged 4.34 times the weekly capacity of the furnaces in blast at the first of the month appears eminently reasonable, for there are precisely 4.34 weeks in the average month.[1] This interpretation of the results of the correlation of weekly capacity with monthly production adds to the confidence with which they may be applied. One might raise the question whether the fact that the monthly production was 4.34 times the weekly capacity from 1903 to 1909 proved that the same relation would hold from 1884 to 1902. Might not conditions be sufficiently different during the two periods to result in a different relationship? Having discovered that this relationship depends on the number of weeks in the average month, we may feel assured that the relationship has not changed.

The fact that the production of pig iron for the average month is equal to the weekly capacity of the furnaces in blast multiplied by the number of weeks in the month, plus a certain constant quantity, suggests the idea that the regression equation can be improved upon by considering the actual length of each month instead of taking each month as of average length. During most years February has only four weeks (in a leap year, 4.14 weeks), while the months containing 30 days have 4.28 weeks and the months containing 31 days have 4.43 weeks. There is no effect from holidays since blast furnaces do not recognize them.

Probably the failure to obtain perfect correlation between weekly capacity and monthly production is due largely to the fact that the monthly production is affected not only by the capacity of the furnaces in blast but also by the length of the month. The one cannot be estimated accurately from the other by means of the simple linear relationship assumed in correlation, hence the correlation is not perfect.

But why should it be found necessary to add 8,100 tons to the figure obtained when the weekly capacity is multiplied by the number of weeks in the month? The answer is easy. Pig iron production showed a gradual increase during this period, so that the average weekly capacity for most months was greater than the weekly capacity at the

[1] Taking 365 days in the year gives 52.14 weeks, or an average of 4.345 weeks to the month.

TABLE II

MEAN BETWEEN WEEKLY CAPACITY OF FURNACES IN BLAST AT THE FIRST OF EACH MONTH AND THE CAPACITY AT THE FIRST OF THE FOLLOWING MONTH, 1884–1902

Month	1884	1885	1886	1887	1888	1889	1890	1891	1892	1893
January	....	60	84	118	112	140	162	145	176	164
February	....	66	85	122	100	138	166	130	179	165
March	....	68	92	124	99	138	168	114	178	169
April	....	67	102	126	104	136	168	105	172	172
May	75	66	108	108	108	130	168	122	165	170
June	74	64	108	94	105	128	166	148	160	156
July	72	64	107	99	105	131	158	160	153	124
August	68	64	104	110	112	132	156	159	144	90
September	65	64	104	125	117	136	162	164	146	73
October	65	66	108	130	124	146	166	172	156	72
November	66	72	110	129	132	154	168	177	164	84
December	62	80	112	124	140	159	163	178	166	94

Month	1894	1895	1896	1897	1898	1899	1900	1901	1902
January	96	161	197	156	222	235	288	260	308
February	101	156	188	162	226	228	288	282	320
March	114	154	183	166	228	232	284	287	324
April	114	153	184	168	228	243	284	292	334
May	82	152	181	166	224	247	288	300	338
June	70	160	176	163	215	253	282	306	340
July	97	172	162	161	206	259	256	300	336
August	128	183	137	171	204	262	230	295	328
September	146	194	114	188	208	267	220	297	333
October	152	204	112	202	216	276	211	307	334
November	161	212	128	215	226	285	214	315	334
December	164	207	146	221	234	288	232	304	342

beginning of the month. The production in January, 1903, was 1,472,000 tons, and in December, 1909, 2,635,000 tons; during 83 months there was an average increase each month of 14,000 tons in the monthly production. This would lead us to expect the production for an average month to be 7,000 tons (one-half of 14,000) greater than would be indicated by the capacity of furnaces in blast at the beginning of the month. The correlation gives 8,100 tons instead of 7,000. Probably the figures on capacity do not apply strictly to the beginning of the month, but to a date averaging about two days before the beginning of the month.

The constant term of 8,100 tons is like the regression coefficient of 4.34 in the equation of regression in that it is a figure which works "on the average," but is not precisely accurate for each month. It appears that for the average month the production was 8,100 tons greater than was indicated by the capacity of furnaces in blast at the beginning of the month, but we know that during some months the production was decreasing so that during those months the total production must have been less than that indicated by the capacity of furnaces in blast at the beginning of the month; there were other months in which the

production was increasing more rapidly than in the average month, so that the total production must have been considerably more than the average of 8,100 tons in excess of the production indicated by the capacity of the furnaces in blast at the beginning of the month.

Probably the best method of adapting the figures on capacity of furnaces in blast at the first of the month so as to give an indication of the monthly production is to take the mean between capacity of furnaces in blast at the first of the month and the capacity of the furnaces in blast at the first of the next month. This figure multiplied by the number of weeks in the month should give a highly reliable indication of the production for the month. The accompanying table shows the weekly capacity, by months, computed by the method developed above. The estimates of monthly production furnish a series which is homogeneous with the statistics available since 1902 and which, like the later figures, shows fluctuations resulting from the variations in the length of months. For some purposes the capacity figures may be

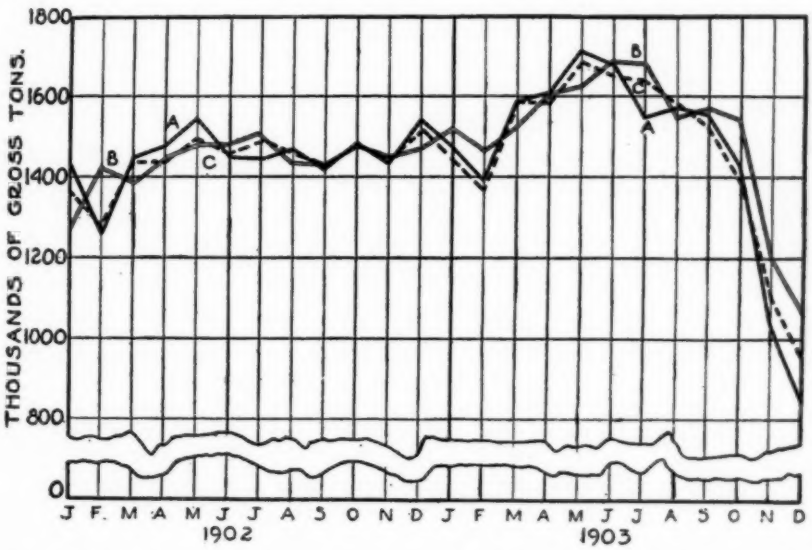


FIG. I. COMPARISON OF RECORDED PRODUCTION OF PIG IRON WITH TWO ESTIMATES BASED ON WEEKLY CAPACITY OF FURNACES IN BLAST THE FIRST OF THE MONTH, 1902–1903.

(A) Recorded production as given by the *Iron Age*.
(B) Miss Myers' estimate based on regression equation.
(C) Estimate by revised method.

The revised method shows a closer correspondence with recorded production through reflecting the effect of varying length of the months and through following the larger fluctuations more accurately.

preferable, inasmuch as they give the true *rate of production* at the time to which they apply.

Figure 1 shows the *Iron Age* data on production for 1902 and 1903 (the heavy solid line) compared with estimates based on weekly capacity of furnaces in blast the first of the month, calculated according to the method just outlined (the dotted line) and according to Miss Myers' method (the hollow line).[1] The chart shows the dotted line following more closely to the solid line than does the hollow line for the two reasons which have been anticipated: It reflects the effect of the varying length of the months and gives a more accurate picture of the changes in periods of sharp upward or downward movements.

The conclusion which follows from the foregoing discussion as it concerns the person who desires monthly data on production of pig iron prior to 1903 is contained in the accompanying table.

TABLE III

MONTHLY PRODUCTION OF PIG IRON, 1884–1902*

Month	1884	1885	1886	1887	1888	1889	1890	1891	1892	1893
January	....	266	372	523	496	620	718	642	780	727
February	....	264	340	488	414	552	664	520	741	660
March	....	301	408	549	439	611	744	505	789	749
April	....	287	437	539	445	582	719	449	736	736
May	332	292	478	478	478	576	744	540	731	753
June	317	274	462	402	449	548	710	633	685	668
July	319	284	474	439	465	580	700	709	678	549
August	301	284	461	487	496	585	691	704	638	399
September	278	274	445	535	501	582	693	702	625	312
October	288	292	478	576	549	647	735	762	691	319
November	282	308	471	552	565	659	719	758	702	360
December	275	354	496	549	620	704	722	789	735	416
Total	....	3,480	5,322	6,117	5,917	7,246	8,559	7,713	8,531	6,642

Month	1894	1895	1896	1897	1898	1899	1900	1901	1902
January	425	713	873	691	983	1,041	1,276	1,152	1,364
February	404	624	778	648	904	912	1,152	1,128	1,280
March	505	682	811	735	1,010	1,028	1,258	1,271	1,435
April	488	655	788	719	976	1,040	1,216	1,250	1,430
May	363	673	802	735	992	1,094	1,276	1,329	1,497
June	300	685	753	698	920	1,083	1,207	1,310	1,455
July	430	762	718	713	913	1,147	1,134	1,329	1,488
August	567	811	607	758	904	1,161	1,019	1,307	1,453
September	625	830	488	805	890	1,143	942	1,271	1,425
October	673	904	496	895	957	1,223	935	1,360	1,480
November	689	907	548	920	967	1,220	916	1,348	1,430
December	727	917	647	979	1,037	1,276	1,028	1,347	1,515
Total	6,196	9,163	8,310	9,296	11,453	13,368	13,359	15,402	17,287

*Calculated by multiplying the figures in Table II by the number of weeks in the month.

[1] Miss Myers' figures, as given in her paper, were taken for 1902 and for 1903, and were calculated from weekly capacity by means of the equation of regression which she gives.

U. S. DIVORCE RATES, 1922

By Walter F. Willcox

The number of divorces per 100,000 married couples in 1922 has been computed for continental United States and the several states on the basis of figures recently issued by the Bureau of the Census in the form of a press bulletin. The number of married couples is assumed to equal the number of married men, this number being in each state greater than the number of married women. The excess, of course, is due to the large number of married immigrants whose wives have not entered the United States. From the number of married couples

TABLE I

States	Divorce rate per 100,000 married couples
District of Columbia	170
New York	186
North Carolina	272
North Dakota	280
New Jersey	293
Georgia	319
Connecticut	339
Wisconsin	367
Pennsylvania	369
Massachusetts	398
South Dakota	437
Delaware	441
Maryland	474
West Virginia	480
Vermont	486
Louisiana	487
Virginia	529
Minnesota	535
Alabama	575
New Hampshire	627
United States	**653**
Rhode Island	654
Mississippi	660
Utah	664
New Mexico	672
Maine	678
Washington	689
Iowa	721
Ohio	749
Illinois	758
Kentucky	807
Tennessee	815
Nebraska	827
Michigan	864
Kansas	878
Idaho	897
Arizona	910
Montana	980
Colorado	990
Florida	1,013
Indiana	1,026
California	1,101
Missouri	1,110
Wyoming	1,144
Texas	1,269
Oklahoma	1,277
Oregon	1,363
Arkansas	1,698
Nevada	6,274

in 1910 and 1920 the number in 1922 has been estimated by the method of arithmetical progression. Table I gives the divorce rate of the states arranged in the order of increasing rate. Table II gives the same facts for the states arranged in geographical order.

Table I shows that of the states in the Atlantic groups only Maine, Rhode Island and Florida have divorce rates above the average for the whole country, and that of the states in the central and western groups only the Dakotas, Minnesota, Wisconsin, Alabama and Louisiana have divorce rates below the average for the whole country. The eastern mountain range thus divides the country into an Atlantic section with a relatively low divorce rate and a central and western section with a relatively high divorce rate. The average divorce rate for the Atlantic states is 340; that of the central and western states is 864, or two and one-half times as great. The difference may be expressed in another way. The average divorce rate of the Atlantic coast states is about one-half (52 per cent) of the average for the whole country. The average divorce rate of the central and western states is one-third (33 per cent) greater than the average of the whole country. If we take the Mississippi River as the dividing line between the east and west, classing Minnesota and Louisiana with the western states, the divorce rate east of the Mississippi is about four-fifths (79 per cent) of the average for the country while that west of the Mississippi is nearly one-half (49 per cent) greater than the country's average.

If attention is confined only to the states east of the Rocky Mountains, the so-called Atlantic and Central Divisions, the divorce rate in the South was greater than that in the North as appears in the following figures:

Northern States (North Atlantic plus North Central)	567
Southern States (South Atlantic plus South Central)	712

The divorce rate in the states west of the Mississippi River exceeds that of the eastern states by nearly nine-tenths; the rate in the southern states exceeds that in the northern by about one-fourth. The extreme rates in the separate states are found in the District of Columbia with the lowest rate, 170, and Nevada with the highest, 6,274, or 37 times as great. But if the Nevada rate be disregarded on the ground that most of the divorces are granted to non-residents, many of them being doubtless invalid, the rate in Arkansas is the highest, about ten times that in the District of Columbia.

Certain states have a rate quite different from what their geographical location might lead one to expect. Thus of the 15 states with a very low divorce rate 12 fall in the Atlantic Division and 3, Wisconsin and the Dakotas, in the Central Division. On the other hand, of the

TABLE II

Division and State	Divorce rate per 100,000 married couples
United States	652
New England	455
Maine	678
New Hampshire	627
Vermont	486
Massachusetts	398
Rhode Island	654
Connecticut	339
Middle Atlantic	272
New York	186
New Jersey	293
Pennsylvania	369
East North Central	768
Ohio	749
Indiana	1,026
Illinois	758
Michigan	864
Wisconsin	367
West North Central	797
Minnesota	535
Iowa	721
Missouri	1,110
North Dakota	280
South Dakota	437
Nebraska	827
Kansas	878
South Atlantic	394
Delaware	441
Maryland	474
District of Columbia	170
Virginia	529
West Virginia	480
North Carolina	272
South Carolina	...
Georgia	319
Florida	1,013
East South Central	721
Kentucky	807
Tennessee	815
Alabama	575
Mississippi	660
West South Central	1,127
Arkansas	1,698
Louisiana	487
Oklahoma	1,277
Texas	1,269
Mountain	1,025
Montana	980
Idaho	897
Wyoming	1,144
Colorado	990
New Mexico	672
Arizona	910
Utah	664
Nevada	6,274
Pacific	1,042
Washington	689
Oregon	1,363
California	1,101

15 states with a very high divorce rate 13 lie west of the Mississippi River, one, Indiana, in the Central states and one, Florida, in the Atlantic group.

SOCIAL SCIENCE ABSTRACTS

Research in social science is hampered by the lack of indexes of abstracts of periodical literature. At the present time our research workers are far more handicapped than chemists or physicists because scientists in these latter subjects have resort to such publications as *Chemical Abstracts* and *Science Abstracts*, wherein are found systematically indexed, careful abstracts covering the whole field. A research worker in these sciences can do his preliminary reading with convenience and expedition. In contrast, the social scientist who wants to exhaust the periodical literature on his subject cannot depend on the existing abstracts published in social science journals, however excellent these may be as individual pieces of work, since the existing abtracts do not fully cover the field. Consequently, he must supplement them by considerable reading from original articles and run the risk of missing something really important.

In view of this situation the present statement has been prepared by the Committee on Social Science Abstracts of the Social Science Research Council. It is hoped that members of the social science societies, as well as any other interested persons, will study this statement, in order that they may be prepared to make suggestions to the chairmen of the committee at any time and particularly to be prepared to take action on one of the alternative plans presented at the annual meetings of the social science societies next December.

At the present time the *American Political Science Review* and the JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION do not print systematic abstracts of their literature. The *American Economic Review* has for some years printed brief abstracts. The *American Journal of Sociology* has since July, 1921, printed in each of its issues several pages of carefully classified abstracts. The original system of classification was enlarged in March, 1922, and since this time has comprised ten main headings and forty-eight sub-headings. Since the abstract plan of the *American Journal of Sociology* is more complete than that of any of the other social science journals, the main headings are presented herewith, with the suggestion that the reader examine the complete classification with its subordinate headings, as it is found in Volume 19, pages 373–374 of the *American Journal of Sociology*. The existing classification is not final, nor have the methods of abstracting been yet perfected, but the scheme is serviceable. The main headings are as follows:

I. Personality: the Individual and the Person.
II. The Family.
III. Peoples and Cultural Groups.
IV. Conflict and Accommodation Groups.
V. Communities and Territorial Groups.
VI. Social Institutions.
VII. Social Science and the Social Process.
VIII. Social Pathology: Personal and Social Disorganization.
IX. Methods of Investigation.
X. General Sociology and Methodology of the Social Sciences.

The following alternative plans for improving the existing abstracts of social science are suggested for consideration by members of the social science societies. It is assumed that only articles containing the results of original research will be abstracted.

(1) Independent and separate publication in bulletin form.

a. Abstract service to be maintained jointly by membership dues from the social science societies, and published as a separate monthly bulletin. (the principles of editorial organization described under (2) c, below are understood to apply to this plan)

or

b. Abstract service to be maintained jointly by social science societies with the cost of administration and publication met in whole or in part by a subsidy or grant of funds from some National Foundation interested in Social Research. (The principles of editorial organization described under (2) c, below are understood to apply to this plan.)

c. Advantages of these plans:

(a) The whole field of social science would be covered with approximate completeness so that cross-referencing would make accessible valuable leads from related fields—on which so much progress depends.

(b) Duplication in abstracting would be avoided, such as would exist if every journal abstracted independently for its own clientele.

(c) One large strong abstract journal could secure by exchange or purchase more currect serial literature than could independent journals, and this would save duplicate subscriptions.

(2) Abstract services of the social science journals continued as at present, but—

a. Present services enlarged to cover the field more intensively and comprehensively.

b. Methodology of abstracting systematized by acceptance of some mutually agreeable plan such as:

c. Acceptance of guidance over abstracting to be exerted by some central editorial body representative of the whole field.

(a) Which formulates a methodology of abstracting to be generally used.

(b) To conform to some objective system of classification of subjects.

(c) Cross referencing.

(d) To avoid unnecessary duplication in printing.

(e) Editing and abstracting paid.

d. Disadvantages of this plan as compared with (1) above:

(a) Danger of duplication of abstracts of the same article in several journals. This is now a real difficulty in the abstracts of biological sciences.

(b) The field would be incompletely covered and cross-referencing incomplete.

(c) Current files of serial literature would not be as complete and there would be inevitable duplication of subscription costs.

(3) Enlargement of the present abstract services and creation of new services in existing social science journals where needed.

a. Organization:

(a) Editor-in-chief, four associate editors, one for each science, but unpaid.

(b) Abstracting paid for by the page.

b. Agreement on a common basis of classification of abstracts, each journal publishing in full all the abstracts relating to the field of its science and printing merely the scheme of classification of each of the other sciences, with a note referring the reader to these journals for full abstracts.

c. The disadvantages of this plan are all the faults of plan (2), but in greater degree.

It will be observed that the last plan is the least ambitious of the three and perhaps in the present stage of development of the social sciences, the most practicable. This last plan would leave to each of the journals the details of its own abstracting, since only a loose type of common editorial organization would exist. On the other hand, the plan has such serious objections that it could be at best but a temporary makeshift.

The cost of the alternative plans would vary and it is probable that the third plan would be the least expensive. In this connection, it may be of interest to note that the abstract editor of the *American Journal of Sociology* now prints galley sheets of the abstracts of each issue. These may be had by all subscribers for $1.00 a year. The subscriber then cuts up the galleys and pastes the separate abstracts on filing cards. The cost of chemical abstracts which is published according to plan(1) a, is $6.80 per subscriber, with a subscribing membership of 13,000. The combined membership of the four social science associations is about 7,000.

F. Stuart Chapin,
Chairman, University of Minnesota.

Members of the Social Science Abstract Committee:

F. Stuart Chapin, Sociology, Minnesota.
A. C. Hanford, Political Science, Harvard.
Davis R. Dewey, Economics, Mass. Institute of Technology.
Walter F. Willcox, Statistics, Cornell.

TRADE STATISTICS

The following correspondence between the Department of Justice and the Department of Commerce of the United States on the collection of statistical data by trade associations shows the status of a matter which is of such keen interest to the business world:

December 11, 1923.

THE HONORABLE,
THE ATTORNEY-GENERAL,
Department of Justice,
Washington, D. C.

My dear Mr. Attorney-General:

The question of the right of trade associations to gather and distribute information and statistics is becoming more and more acute. My own impression is that the collection and distribution of current unidentified information and statistics as to production, stocks on hand, and prices on closed transactions should be permissible providing that it is all made available on fair terms, not only to the association members, but to all others interested, including the general public.

Such knowledge is essential to the consumer, producer, manufacturer and distributor, as it places them all in a position to interpret and judge market conditions intelligently on the basis of supply, demand and current prices, and to gauge their purchases and sales accordingly. This unquestionably would be of tremendous aid in the economical conduct of business and redound to the public good.

If business be compelled to operate without such vital information it will naturally be forced into unscientific and highly speculative avenues.

Competition based on fair and equal information of existing conditions would more likely result in lower prices to the consumer than competition based on uncertainty, in which each dealer must add something to his price to cover unforeseen eventualities.

Some time ago I realized that the carrying out of the purposes of this Department as set forth in the Organic Act, "To foster, promote, and develop a foreign and domestic commerce, the mining, manufacturing, shipping and fishing industries, and the transportation facilities of the United States," required that the character of information described in the first paragraph hereof should be available, and that the gathering of such from the individual units of industry would involve such a gigantic physical task that this Department with its facilities could not undertake such action with the faintest hope of attaining the ends desired. I therefore have in the past utilized to a very considerable extent the trade association as a medium for securing such information and have received splendid coöperation from them as a whole.

Seeking to clarify the situation regarding legitimate trade association activities, I set forth my views in several letters to you in February, 1922, and requested your informal opinion as to the legality of many association functions, including the collection and distribution of the character of information herein referred to. Predicated principally upon your informal views in reply thereto, this Department formulated a plan of coöperation with trade associations (copy of which is attached hereto) under which the association's secretary collects and compiles unidentified current information and distributes it to the members simultaneously sending identical reports to governmental agencies, competitors of the association's members, and to any other person who arranges for them. Under this plan no supplemental or separate reports are transmitted to the members only. This Department receives all such reports for wide dissemination by publication.

My attention has been directed to a decree entered November 26, 1923, in the District Court of the United States for the Southern District of Ohio, in the case of United States *vs.* Tile Manufacturers Credit Association, et al., paragraph two of page four of which is as follows:

> "Provided, however, that the defendants are not restrained or enjoined from maintaining an association, either voluntary or incorporated, for the following objects and purposes and none other,"

after which various permissible activities are set forth. Paragraph two of page three is as follows:

> "Provided, however, that the defendants may, through the association, or corporation hereinafter provided for, receive and compile for transmission to any governmental agency such information and statistics as it may request as to the production, shipments, the stocks on hand and the prices of tiles, but are restrained from distributing said information among themselves, except that information respecting sales may be collected annually and used to enable the assessment of the several members for their proportionate parts of the several expenses of the association, and for no other purpose."

I interpret the last quoted paragraph to mean that it would be unlawful for the proposed association on behalf of its members to transmit information and statistics of the character therein described to its members, but that it could receive and compile it for the purpose only of transmitting it to a governmental agency that might so request.

It is my understanding that this decree is only binding between the parties thereto; however, in view of informal conversation between representatives of your Department and this Department on the subject of trade association activities, I am inclined to be of the opinion that the last mentioned paragraph embodies an expression of the present policy of your Department relative to the collection, compilation and distribution of information and statistics of the character therein set out applicable to trade associations in general.

It is not the desire or purpose of this Department to continue operations under the coöperative plan if it is in conflict with the policy of your Department; it is our desire, however, to call your attention to the situation that in my opinion will develop, if my interpretation of this decree correctly expresses the policy of your Department. I think there is great likelihood that not only the associations from which this Department now receives valuable statistics but a great many others, will discontinue the collection of information and statistics as to production, shipments, stocks on hand, and the prices on closed transactions. They will not go to the expense of collection, if the only use that can be lawfully made of them is to transmit them to some governmental department. If this should happen, I fear that the efficiency of this Department in carrying out the purposes set forth in the act creating it would be very greatly impaired.

I respectfully request that you informally advise me, in view of the foregoing, whether or not this Department should discontinue its present plan of coöperation with trade associations.

Yours faithfully,

HERBERT HOOVER,
Secretary of Commerce.

OFFICE OF THE ATTORNEY-GENERAL
WASHINGTON, D. C.

December 19, 1923.

My dear Mr. Secretary:

Your communication of the 11th instant relating to exchange of statistics through trade associations was received, and has been given careful consideration; and in reply thereto I beg to say:

Referring to the correspondence which passed between your Department and this Department in February, 1922, I call your attention to the following paragraphs in the initial communication, transmitted by you to this Department on February 3, which contain the sole reference in that letter to the gathering of statistics by trade associations and their dissemination:

> "(10)–A. May a trade association collect statistics from each member showing his volume of production, his capacity to produce, the wages paid, the consumption of his product in domestic or foreign trade, and his distribution thereof, specifying the volume of distribution by districts, together with his stock, wholesale or retail?
>
> "B. And may such trade association, on receipt of the individual reports of each member, compile the information in each report into a consolidated statement which shows the total volume of production of the membership, its capacity to produce by districts of production, which, in some instances, include a state or less area, the wages by districts of production, the consumption in foreign or domestic trade by districts, the volume of distribution by districts, and the stocks on hand, wholesale and retail, by districts?
>
> "C. And, if, after compiling the information as aforesaid, the information received from members as well as the combined information is not given by the association to any other person, may it then file the combined statement with the Secretary of Commerce for distribution by him to the members of the association through the public press or otherwise and to the public generally and to all persons who may be in any way interested in the product of the industry, it being understood that the individual reports for the members should cover either weekly, monthly, quarterly, or longer periods as may be deemed desirable by the

members, and, when a period is adopted, the report for each member shall cover that period, and the combined report shall be for that period?

"(11)–A. May a trade association, at the time it collects the production and distribution statistics above outlined, at the same time have their members report the prices they have received for the products they have sold during the period taken, specifying the volume of each grade, brand, size, style, or quality, as the case may be, and the price received for the volume so sold in each of the respective districts where the product is sold?

"B. And may the association, without making known to any person the individual price reports of any member, consolidate all of the reports into one, and show the average price received for the total volume of each grade, brand, size, style, or quality, as the case may be, distributed in each district covered by the distribution statistics for the period covered by each individual report?

"C. And may the association, after making such compilation, send the compiled report as to average price, as aforesaid, to the Secretary of Commerce, to be by him distributed to the public and to any or all persons who may be interested in the particular industry making the reports?"

This particular subject had been carefully considered by representatives from both Departments before the foregoing was incorporated in the communication; and my views as to how far trade associations should be used in collecting and distributing statistical information are there set forth.

The paragraph to which you refer in the decree entered in United States *vs.* Tile. Manufacturers Credit Association reads as follows:

"Provided, however, that the defendants may, through the association, or corporation hereinafter provided for, receive and compile for transmission to any governmental agency such information and statistics as it may request as to the production, shipments, the stocks on hand and the prices of tiles, but are restrained from distributing said information among themselves, except that information respecting sales may be collected annually and used to enable the assessment of the several members for their proportionate parts of the several expenses of the association, and for no other purpose."

You will observe that this proviso in the decree complies strictly with the paragraphs above quoted from your letter of February 3, 1922.

Two objects were had in mind during the conferences which preceded the above mentioned correspondence, both of which it was thought were secured by the limitations embraced in the above quoted paragraphs of your letter: First, that the information distributed should be general, and second, that individual contact between those engaged in the same industry with reference to matters which vitally affect prices should be avoided. If the character of the information and the manner of its dissemination be restricted as specified in your letter, one member would not be informed as to the individual activities of another member, and those engaged in the industry would be prevented from revealing their business to their competitors. Under the system now practiced by many of the associations each member reports its production, shipments, stocks on hand, and each individual sale, stating the price at which it is made and, generally, the locality where made; and this information is distributed by the secretary or manager of the association among all the members, though in some associations the names of the members making the sales are omitted. Thus each member reveals the details of his entire business to every other member, which, as suggested by the Supreme Court in the Hardwood case, is entirely inconsistent with the normal attitude of real competitors.

In my judgment the effect of general information as to the conditions of an industry, such as the total production, shipments, stocks on hand and the average price, or range of prices, is entirely different from that resulting from each person engaged in an industry receiving directly, or through a common medium, reports which reveal to him the exact condition of the business of all of his competitors. When thus informed each one is invited, and is naturally inclined, to imitate the conduct of his most successful competitor; and the spirit of comradeship created by the confidential exchange of information of this character necessarily prevents the free competition between them which would otherwise prevail.

Those who organize and conduct these associations appear to entertain the idea that if the information imparted relates only to past and closed transactions there can

be no violation of the Anti-Trust Act. In my judgment such an idea is wholly fallacious. One's future conduct is to be judged by what he has done and is then doing, and not so much by what he says he will do. It is one's actual conduct that is taken as an example for imitation. It has developed in the trial of cases involving associations that the members first agreed upon prices; but such a plan did not work because the members could not be relied upon to keep the agreement; and the system of exchanging statistics was adopted because it was found to be the only effective way to procure coöperation as to prices and production; and such coöperation could thus be procured even in the absence of any positive agreement.

Again, the idea seems to be prevalent that no exchange of information between the members, regardless of its extent or character, can be unlawful if at the same time publicity be given thereto through the press or some governmental agency. In my judgment this idea is likewise fallacious. The illegality as well as the evil results arise from the *coöperation among the members pursuant to a positive or tacit understanding;* and this coöperation is not affected by publicity. Those who purchase the commodity, though fully informed as to the activities of the association, can protect themselves only by an organization and coöperation of like character, which, if it were lawful, is an impossibility upon the part of the public.

I have no doubt that it is important that those engaged in an industry have general information as to the conditions of that industry, but I think that information should be distributed strictly through a responsible medium, like your Department; and I see no objection to its being gathered by an association provided it be strictly guarded and the association be prohibited from distributing it among its membership. This is the same view that I entertained when the communications were exchanged in February, 1922; and it has since been strongly confirmed by decisions of the Supreme Court, and by investigations of a number of associations and the trial of cases involving associations.

This is but a statement of the position I feel impelled to take as Attorney-General of the United States in enforcing the Anti-Trust Act. But, of course, as to what activities and how far you will coöperate with trade associations are matters for your determination in conducting your Department.

Yours sincerely,

H. M. Daugherty,
Attorney-General.

Honorable Herbert Hoover,
Secretary of Commerce,
Washington, D. C.

PROGRESS OF WORK IN THE CENSUS BUREAU

QUINQUENNIAL CENSUS OF AGRICULTURE

Congress appropriated $3,500,000 for the census of agriculture to be taken in the year 1925 as authorized by the Act of 1919 providing for the Fourteenth Census. This is half a million less than the amount asked for, and necessitates therefore some curtailment in the plans for the undertaking. Preparations for this census are now in progress. The schedule, which was formulated in consultation with representatives of the Department of Agriculture and the Census Advisory Committee, has been sent to the printer for an edition of 15,000,000, that number being necessary to provide for the six and a half million or more farms to be canvassed, as large numbers of extra schedules will be mailed to farmers, farm papers, agricultural colleges, etc., in advance of the enumeration. This schedule, although much less detailed than the one used for the Fourteenth Census, is comprehensive, covering acreage of farm land and of crops, value of farm property, live stock, farm products, etc. The agricultural schedule is always a rather elaborate affair, mainly because of the large variety of crops and the

various classes of domestic animals which have to be separately reported. The present schedule carries 176 separate inquiries, which represents a considerable reduction as compared with the 400 or more items on the schedule of 1920. It repeats the most important of the inquiries which were made in 1920, and contains also a few new questions designed to throw as much light as possible on phases of agriculture which are now undergoing significant changes.

As has been customary the schedule in asking for farm acreage distinguishes the main classes of farm land. But the traditional distinction between improved and unimproved land has not been retained. The definition of improved land is necessarily vague or elastic, and it has been differently interpreted in different parts of the country. The new questions call specifically for the acreage of crops harvested in 1924, land from which no crop was harvested because of crop failure, land in cultivated summer fallow, crop land lying idle, different kinds of pasture, etc.,—in all nine classes of land. These questions are designed to show the present tendencies in the utilization of the farmland resources of the country.

Among the new questions on the schedule are two with regard to the amount of taxes paid by farmers in 1924. The farmers' burden of taxes, particularly state and local taxes, has been increasing very rapidly of late, but there are no figures available to show the amount of these taxes. Another new question, asking for the kind of road which adjoins the farm, is somewhat related to the question on taxation, since a considerable part of the local taxes are expended for the building of roads.

There is, as usual, an inquiry with regard to farm mortgage debt, and to this a question has been added asking for the amount of other debts owed by the farmer, that is, debts not secured by real estate mortgage.

THE FIRST BIENNIAL CENSUS OF MANUFACTURES

The final and complete report presenting the results of the 1921 census of manufactures is now being printed. This is an event of more than ordinary interest as it marks the completion and permanent publication of the first biennial census of manufactures. It may seem that this is a late date at which to be printing the results of a census covering the year 1921, but practically all the statistics contained in this volume have been previously published in one form or another. The collection of the data for any census of manufacturing industries cannot of course be begun until a few weeks after the year to which the census relates. The canvass for the census of 1921 was begun about the first of February, 1922, and practically completed by the first of July. Returns were received from 250,266 establishments. The tabulation of the data was started in July, 1922, and completed by the middle of June, 1923. The first publication of results was a press summary relating to the carpet and rug industry, issued on July 31, 1922. Between that date and May 12, 1923, 287 press summaries covering that many industries were issued. The first printed bulletin, covering the rice cleaning and polishing industry, was published May 24, 1923, and the last, that on the wearing apparel industry, on March 4, 1924. This is looked upon by the Bureau as a fairly creditable record in point of time con-

sumed although it is believed that it can be considerably shortened in the future, as the manufacturers become acquainted with the value of the statistics and more accustomed to preparing the biennial reports.

The final report of the census of 1921 is a single volume of 1,637 pages octavo size. The first part of the volume, in which the figures are presented by industries, reprints the 56 industry bulletins previously published and incorporates also the data for the minor industries previously published in press summaries only. The last part contains summaries for states and cities by industries, chapters on monthly employment of wage earners, hours of labor, size of establishment, output in relation to capacity, and general tables.

The total cost of this biennial census of manufactures exclusive of printing was $741,738, of which $310,394 was for field work and $431,344 for work in the office. The cost of printing was $35,840 of which $14,011 was for job work (schedules, forms, stationery, etc.) and $21,829 was for the printed bulletins and final report.

In the meantime the work of taking the second biennial census of manufactures, covering the year 1923, is well under way, the schedules having been collected and the tabulation begun, and 32 press summaries on various industries having been issued thus far.

ESTIMATING POPULATION

The problem of estimating the population of states and cities in intercensal years so as to provide a basis for rates and ratios such as birth rates, death rates, per capita taxes, etc., has never been satisfactorily solved, and is perhaps giving more difficulty at present than is usual because of the abnormal conditions of population growth that prevailed in the years immediately preceding the census of 1920. The matter was recently called to the attention of the Census Advisory Committee and by them referred to a subcommittee consisting originally of Walter F. Willcox and Robert E. Chaddock, and later enlarged by the inclusion of Willford I. King and Robert C. Woodbury.

The practice of making population estimates was established in 1903, or practically as soon as the Bureau was made a permanent office charged with the continuous collection of statistics. The matter was given careful consideration at that time by a committee of statisticians who reached the conclusion that the best practicable method of estimating, all things considered, was the so-called arithmetical method, which assumes that the population of a given state or city increases by a uniform amount each year, this amount being taken as equal to the average annual increase in the interval between the last two censuses. Tests that were made indicated that in the great majority of cases this assumption of a constant amount of increase gave results sufficiently accurate for statistical purposes, and came nearer the truth than the alternative assumption of a constant rate of increase. But while this method of estimating, which was adopted in 1903 and has been followed ever since, seemed to be the best practicable method its results have not always been satisfactory. It fails when there are radical changes or abnormal conditions in population growth such as a sudden expansion of the population of a city, resulting perhaps from the

introduction of new industries, or a sudden check to a city's growth following the collapse of a boom.

For the United States as a whole, a better estimate than that obtained by the arithmetical method can doubtless be arrived at by utilizing the statistics of births and deaths and immigration and emigration. But unfortunately that method cannot be used for individual states and cities, as we have no data as to the migration of population across state lines or city boundaries. The most satisfactory solution of the problem would be a quinquennial census of population. For the first few years after a census is taken almost any method of estimating will work well in the great majority of cases; but the farther away we get from the census the more unsatisfactory are the results. And so long as we have only a decennial census, estimates for the second half of the decade will probably be unreliable and uncertain at best.

COMMITTEE ON FINANCIAL STATISTICS OF CITIES

The Director of the Census attended the Nineteenth Annual Convention of the National Association of Comptrollers and Accounting Officers held in Providence, Rhode Island, June 17, 18, and 19 and delivered an address on "Financial Statistics of Cities."

He emphasized the importance of coöperation between the city officials and this Bureau; also the importance of a more extended adoption of a uniform system of classification of accounts; and the desirability of having annual reports of cities gotten out more promptly.

As a result of this discussion, a resolution was passed providing for the appointment of a permanent committee to work in close harmony with the Director and other officials of the Census Bureau in regard to the form of schedules to be used in collecting statistics of the financial transactions of cities and the form of publication carrying these statistics, and to devise methods for the officials of the different cities to coöperate more closely with the Bureau in compiling statistics annually, the object being to publish the data more promptly, secure greater accuracy, and arrange for the cities to bear a portion of the expense of assembling these data. The committee appointed consists of C. C. Pashby, City Clerk, Memphis, Tenn.; Walter R. Darby, State Commissioner of Municipal Accounts, Trenton, N. J.; and W. S. Hanna, Deputy Comptroller, Baltimore, Md.

J. A. H.

WORK OF THE ASSOCIATION'S COMMITTEE ON PROFESSIONAL AND EDUCATIONAL STANDARDS FOR STATISTICIANS

The Committee outlined for its first task the preparation of a list of the available standard texts on statistics. A preliminary selected list of statistical texts is now ready for the membership. It was thought that the publication of this list would be of value to teachers of statistics in guiding them in the collection of suitable material for teaching purposes and that it would be of use, also, to individual students. In the autumn, the Committee expects, also, to conduct a

survey of statistical teaching in American universities and colleges. Members of the Association engaged in statistical teaching in the United States are invited to assist the Committee in the proposed survey. The aim of this survey will be to determine, first, the prerequisites in mathematics and other subjects for students entering statistical classes, the departments of the universities or colleges in which statistical courses are given, the texts used, the extent of laboratory work, the nature of laboratory equipment, and other pertinent facts.

AMERICAN STATISTICAL ASSOCIATION
COMMITTEE ON PROFESSIONAL AND EDUCATIONAL
STANDARDS FOR STATISTICIANS
JAMES W. GLOVER, *Chairman*

SELECTED LIST OF TEXTS IN STATISTICS

Prepared by the Committee on Professional and Educational Standards of the American Statistical Association

I. GENERAL

BAILEY, WILLIAM BACON AND JOHN CUMMINGS
Statistics. Chicago: McClurg, 1917. 153 pp. $.60.

BALLOD, CARL
Grundriss der Statistik enthaltend Bevoelkerungs—Wirtschafts—Finanz—und Handels—Statistik. Berlin: Guttentag, 1913. 348 pp. New York: Stechert. $2.60.

BOWLEY, ARTHUR LYON
Elements of statistics. 2 V. 4th Ed. New York: Scribner, 1921. 474 pp. $8.00. Separately, V. 1, $5; V. 2, $4.
Elementary manual of statistics. New York: Scribner, 3rd Ed., 1920. 220 pp. $3.50.

ELDERTON, WILLIAM PALIN AND E. M. ELDERTON
Primer of Statistics. New York: Macmillan, 1910. 86 pp. $.60.

KING, WILLFORD ISABELL
Elements of statistical method. New York: Macmillan, 1912. 250 pp. $1.50.

KOREN, JOHN, Ed.
History of statistics—their development and progress in many countries. American Statistical Association. New York: Macmillan, 1918. 773 pp. $7.50.

SECRIST, HORACE
Readings and problems in statistical methods. New York: Macmillan, 1920. 427 pp. $2.75.
Introduction to statistical methods. New York: Macmillan, 1917. 482 pp. $2.

ZIZEK, FRANZ
Grundriss der Statistik. Munich and Leipzig: Duncker u. Humblot, 2nd Ed., 1923. New York: Stechert. $4.50.
Statistical averages: a methodical study; auth. tr., with additional notes and references by Warren Milton Persons. New York: Holt, 1913. 392 pp. $2.50.

II. MATHEMATICAL

BACHELOR, LOUIS
Calcul des Probabilités. Paris: Gauthier-Villars, 1912. V. 1, 516 pp. New York: Stechert. $6.

BLASCHKE, E.
Vorlesungen uber Mathematische Statistik. Leipzig u. Berlin: Teubner, 1906. 268 pp. New York: Stechert. $1.85. (Out of print.)

BOREL, EMILE
Elements de la Theorie des Probabilités. 3d Ed. Paris: J. Hermann, 1924. 226 pp. New York: Stechert. $2.30.

BRUNT, DAVID
Combination of observations. New York: Putnam, 1917. 219 pp. $2.50.

CHARLIER, C. V. L.
Vorlesungen uber die Grundzuge der Mathematischen Statistik. Lund, Verlag Scientia, 1920. 125 pp. New York: Stechert. $3.

CZUBER, EMANUEL
Wahrscheinlichkeitsr echnung; und ihre Anwendung auf Fehlerausgleichnung; Statistik und Lebansversicherung. 2 V. 3rd Ed. Leipzig u. Berlin: Teubner. V. 1, 4th Ed., 1924, $3; V. 2, 3rd Ed., 1921, $3.50.

ELDERTON, WILLIAM PALIN
Frequency curves and correlation. New York: Spectator, 1903. 172 pp. $3.50.

FISHER, ARNE
Mathematical theory of probabilities and its application to frequency curves and statistical methods. V. 1, 2nd Ed. New York: Macmillan, 1922. 239 pp. $5.

FORSYTH, C. H.
Introduction to the mathematical analysis of statistics. New York: Wiley, 1924. 241 pp. $2.25.

GLOVER, JAMES WATERMAN, Ed.
Tables of applied mathematics in finance, insurance, statistics, compound interest functions and logarithms of compound interest functions, life insurance and disability insurance functions, probabliity and statistical functions and seven place logarithms of numbers from 1 to 100,000. Ann Arbor: Wahr, 1923. 676 pp. $4.50.

GRIFFIN, FRANK LOXLEY
Introduction to mathematical analysis. New York: Houghton, 1921. 512 pp. $2.75.

JONES, DAVID CARADOG
First course in statistics. Chicago: Open Court Publishing Co., 1921. 286 pp. $3.75.

KELLEY, TRUMAN L.
Statistical method. New York: Macmillan, 1923. 390 pp. $4.

LAPLACE
Oeuvres completes de Laplace. Tome VII. *Theorie Analytique des Probabilites.* Paris: Ganthier-Villars, 1886. 832 pp. New York: Stechert. $8–$9.

MACMAHON, PERCY ALEXANDER
Introduction to combinatory analysis. New York: Macmillan, 1921. 71 pp. $2.60.

MORTARA, GIORGIO
Elementi di Statistica. Rome: Athenaeum, 1917.

PEARSON, KARL
Tables for statisticians and biometricians. Univ. of Chicago Press, 1914. 143 pp. $2.50.

RIETZ, H. L., Ed.
Handbook of mathematical statistics. Prepared under the auspices of the National Research Council. New York: Houghton, 1924. $4.

RUGG, HAROLD ORDWAY
Statistical methods applied to education. New York: Houghton, 1917. 410 pp. $2.

RUNNING, THEO. RUDOLPH
Empirical formulas. New York: Wiley, 1917. 144 pp. $1.40.

SOPER, H. E.
Frequency arrays. New York: Macmillan, 1923. 48 pp. $1.25.

THIELE, T. N.
Theory of observations. New York: Spectator, 1903. 143 pp. $6.

URBAN, F. M.
Grundlagen der Wahrscheinlichkeitsrechnung und der Theorie der Beobachtungsfehler. 1923. 274 pp. New York: Stechert. $2.

VON SANDEN, HORST
Practical mathematical analysis. New York: Dutton, 1924. 195 pp. $4.50.

WEST, CARL JOSEPH
Introduction to mathematical statistics. Adams, R. G., 1918. 150 pp. $2.50.

WELD, LEROY DOUHERTY
Theory of errors and least squares. New York: Macmillan, 1916. 190 pp. $1.25.

WHITTAKER, E. T. AND G. ROBINSON
Calculus of observations. London: Blackie and Son, 1924. 398 pp. 18 *s.*

WICKSELL, SVEN D.
Elementen av Statistikens Teori. Lund, 1920. 176 pp.

WOODWARD, ROBERT SIMPSON
Probability and theory of errors. 4th Ed. New York: Wiley, 1906. 49 pp. $1.25.

YULE, G. UDNY
An introduction to the theory of statistics. 5th Ed., enlarged. London: Griffin, 1919. 398 pp. $4.25.

III. VITAL STATISTICS

ACTUARIAL SOCIETY OF AMERICA
American-Canadian mortality investigation; based on the experience of life insurance companies of the U. S. and Canada during the years 1900 to 1915, inclusive, on policies issued from 1843 to 1914, inclusive. 2 V., ea. $10. 1918–1919. New York: Actuarial Society of America, 256 Broadway.

ASSOCIATION OF LIFE INSURANCE MEDICAL DIRECTORS OF AMERICA. *Medico-actuarial mortality investigation.* 5 V. $42.50. 1912–1915. Actuarial Society of America, 256 Broadway.

BURN, JOSEPH
Vital statistics explained. London: Constable & Co., 1914. 140 pp.

CZUBER, EMANUEL
Mathematische Bevolkerungstheorie. Leipzig u. Berlin: Teubner, 1923. 357 pp. New York: Stechert. $3.

ELDERTON, W. PALIN AND RICHARD C. FIPPARD
Construction of mortality and sickness tables. London: Black, 1914. 120 pp. (New York: Baker and Taylor. $.67.)

FISHER, ARNE
Elementary treatise on frequency curves and their applications in the analysis of death curves and life tables. New York: Macmillan, 1923. 240 pp. $4.

GLOVER, JAMES W.
United States Life Tables, 1890, 1901, 1910 and 1901–1910. 496 pp. Washington, D. C.: Supt. of Documents, Gov't Printing Office, 1921. $1.25.

HENDERSON, ROBERT
Mortality laws and statistics. New York: Wiley, 1915. 111 pp. $1.25.

HENDERSON, ROBERT AND H. N. SHEPPARD
Graduation of mortality and other tables. (Actuarial Studies No. 4.) New York: Actuarial Society of America, 256 Broadway, 1919. 82 pp. $2.

MOIR, HENRY AND OTHERS
Sources and characteristics of the principal mortality tables. (Actuarial Study's No. 1.) New York: Actuarial Society of America, 256 Broadway, 1919. $2.

MURPHY, RAY D. AND PERCY C. H. PAPPS
Construction of mortality tables from the records of insured lives. (Actuarial Studies No. 2.) New York: Actuarial Society of America, 256 Broadway, 1922. 72 pp. $2.

NEWSHOLME, ARTHUR
The elements of vital statistics. New York: Appleton, 1923. 623 pp. $7.

PEARL, RAYMOND
Medical biometry and statistics. Philadelphia: Saunders, 1923. 379 pp. $5.

TRASK, J. W.
Vital statistics. Washington, D. C.: Supt. of Documents, Government Printing Office, 1915. Pa., $.10.

WHIPPLE, G. C.
Vital statistics. 2nd Ed. New York: Wiley, 1923. 579 pp. $4.

IV. ECONOMICS

CRUM, W. L. AND A. C. PATTON
Introduction to the methods of economic statistics. New Haven: Harty-Musch Press, Inc. Part 1, 1923, 43 pp. Part 2, 1924 44–95 pp.

HARVARD UNIVERSITY COMMITTEE ON ECONOMIC RESEARCH
Review of economic statistics and supplements. Preliminary V. 1. Cambridge Harvard University Press, 1919.

MAYR, GEORG. VON
Statistik und Gesellschaftslehre. Freiburg u. Tubingen: Mohr, V. 1, 2nd Ed., 1914, $4; V. 2, 1897, 2nd Ed. in press; V. 3, Part 1, 1909; V. 3, 1917, $8.25. New York: Stechert.

SMITH, RICHMOND MAYO
Statistics and sociology. New York: Lemcke, 1910. 399 pp. $3.
Statistics and economics. New York: Macmillan, 1899. 467 pp. $3.

WALKER, FRANCIS AMASA
Discussions in economics and statistics. 2 V. New York: Holt, 1899. $6.

V. BUSINESS

BABSON, ROGER WARD
Business barometers used in the accumulation of money; a text-book on applied economics for merchants, bankers and investors. 11 Ed Wellesley Hills: Babson Institute, 1921. 435 pp. $2.50.

BODDINGTON, A. LESTER
Statistics and their application to commerce. London: Lynch, 1921. 220 pp.

COPELAND, MELVIN THOMAS
Business statistics. Cambridge: Harvard Univ. Press, 1917. 696 pp. $3.75.
DUNCAN, CARSON SAMUEL
Commercial research: an outline of working principles. New York: Macmillan, 1919. 385 pp. $2.25.
FISHER, IRVING
Making of index numbers. New York: Houghton, 1923. 560 pp. $7.50.
FREDERICK, JUSTUS GEORGE
Business research and statistics. New York: Appleton, 1920. 342 pp. $2.50.
MILLS, FREDERICK C.
Statistical methods applied to economics and business. New York: Holt, 1924. (In press.)
SECRIST, HORACE
Statistics in business; their analysis, charting and use. New York: McGraw, 1920. 137 pp. $1.75.

VI. GRAPHICS

BRINTON, WILLARD COPE
Graphic methods for presenting facts. New York: Eng. Mag., 1914. 371 pp. $4.
GILMAN, STEPHEN
Graphic charts for the business man. Chicago: La Salle Extension University, 1921. 62 pp. $.60.
HASKELL, ALLEN CECIL AND J. G. BREAZNELL
Graphic charts in business; how to make and use them. Codex Bk, 1922. 250 pp. $4.
KARSTEN, KARL G.
Charts and graphs; an introduction to graphic methods in the control and analysis of statistics. New York: Prentice-Hall, 1923. 724 pp. $6.
LIPKA, JOSEPH
Graphical and mechanical computation. 2 V. New York: Wiley, 1921. 122, 145 pp. Ea. $2.
MARSHALL, WILLIAM CROSBY
Graphical methods for schools, colleges, statisticians, engineers and executives. 1st Ed. New York: McGraw, 1921. 253 pp. $3.

June 10, 1924

MISCELLANEOUS NOTES

Company Houses in Massachusetts Cotton Textile Centers.—The Massachusetts Department of Labor and Industries has made a report on "Company Houses in Massachusetts Owned or Controlled by Cotton Textile Manufacturers." The inquiry was supplementary to the "Special Investigation into Conditions in the Textile Industry in Massachusetts and the Southern States," made by the department.

In the 36 municipalities studied, 62 housing communities were reported, comprising 2,564 houses, containing 31,533 rooms. The number of persons normally employed in the establishments represented was 46,772, of which number 10,257, or 21.9 per cent, occupied company houses. A grouping of the cities and towns according to population discloses the interesting fact that in the more populous cities where other than company houses are available for occupancy by employees, a relatively small

percentage of the number of cotton textile operatives occupy company houses. For example, in the three cities having a population of over 100,000 (Group A) only 4.5 per cent of the employees represented occupied company houses, and for cities in Group B (50,000–100,000) the percentage was only 7.6, whereas in towns in Group E (5,000–10,000), 43.7 per cent of the operatives occupied company houses and in towns in Group F (less than 5,000) 40.6 per cent were so housed.

The discussion of the information obtained also had reference to the type of construction and materials used and rentals charged. Inquiry was made as to utilities furnished and the extent to which lawns and garden plots were supplied with dwellings. In 60 of the 62 communities water is supplied. Gas and electricity are not generally installed in company houses. Only eight of the 62 communities lacked sewer connection. Bathrooms are not standard equipment, although they are provided in a majority of the communities. Lawns and garden plots are not planned for in many communities, but no objection to either is made, should the occupants desire to arrange and care for them.

In general, the inquiry disclosed the fact that in the large mill cities in this state the practice of housing operatives has, in considerable measure, been abandoned. Although no specific inquiry was made on this point, it appears that the provision of company houses in the less populous communities is primarily for the purpose of increasing the supply of houses for operatives, where otherwise a shortage might exist, rather than to provide accommodations in company houses for *all* employees who might wish to reside therein. In the southern states many of the mills recently erected have been located in rural communities, and it has been necessary to provide accommodations for a very large percentage of the operatives in those mills and the mill owners have found it necessary or advisable to maintain the villages and, in many instances, to support in whole or in part the public schools and expend large sums toward the maintenance of churches and welfare work of various kinds. In the northern mill centers there has been a growing disposition on the part of the operatives to select their place of abode and independently make provision for their household and social requirements.

R. F. P.

An International Mathematical Congress was held in Toronto from Monday, August 11, to Saturday, August 16, 1924, under the auspices of the University of Toronto and the Royal Canadian Institute. Section V of this congress dealt with statistics, actuarial science and economics.

The Federal Reserve Bulletin for July, 1924, contains a mid-year review of banking statistics and an account of the statistical work of the Federal Reserve Board during the past year. The custom of presenting retrospective data on banking statistics in the July issue has been in operation for several years. Regular tables, particularly those dealing with foreign industrial and financial conditions, also include data for a number of earlier months.

The Annual Report of the Federal Reserve Board for 1923 contains, in addition to data for the year, a number of tables showing statistics for the Federal reserve banks, member banks, and non-member banks since the organization of the Federal Reserve System. It also, for the first time, includes a table giving all the Federal Reserve Board's index numbers for the entire period for which they are available.

The Industrial Arts Index, a guide to current publications of technical, industrial and financial information, has begun indexing the JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION.

ISAAC HOURWICH

Dr. Isaac A. Hourwich, who has been a member of the Statistical Association since 1909, died on July 23, at the age of 64. Dr. Hourwich was at one time professor of economics in Columbia University, later he was translator at the Bureau of the Mint, and in 1903 he became a statistician at the Census Bureau. Recently he has served as counsel for the Soviet Government Bureau in New York.

DUDLEY A. SARGENT

Dr. Dudley A. Sargent, the head of the Sargent School, died at Peterboro, New Hampshire, on July 21, at the age of 75. He was a graduate of Bowdoin College and received his M.D. from Yale in 1878. For forty years he was director of physical education at Harvard University, and it was in connection with this office that he developed most of the modern gymnasium apparatus used in this country. He has been a member of the American Statistical Association since 1892.

Dr. Robert M. Woodbury has resigned from the United States Department of Labor as Statistician for the Children's Bureau and has accepted an appointment on the staff of the Institute of Economics.

Dr. Raymond Pearl, Director of the School of Hygiene and Public Health, Johns Hopkins University, has returned to the United States after a summer in Europe where he attended various scientific meetings.

Dr. Paul H. Douglas has been granted leave of absence from his Associate Professorship at the University of Chicago for the academic year, 1924–25, which he will spend at Amherst College as Visiting Professor of Political Economy.

Shelby M. Harrison has been appointed Vice-General Director of the Russell Sage Foundation. He will continue for the present as Director of the Department of Surveys and Exhibits of the Foundation, and in charge of the Studies of Social Conditions of the Regional Plan of New York.

Miss Jessamine S. Whitney of the National Tuberculosis Association has been granted a leave of absence to serve as Statistical Advisor for the three demonstrations of the Milbank Memorial Fund.

Dr. Marcus Nadler has left the Division of Research and Statistics of the Federal Reserve Board to accept a position with the National Bank of Commerce in New York City.

Mr. H. A. Widenman, who has recently returned to this country after two years of banking experience in Germany, was engaged in research work for the Division of Research and Statistics of the Federal Reserve Board during the summer.

The publisher, George Wahr, of *Tables of Applied Mathematics*, by James W. Glover, announces that Part III, which includes the tables of probability and statistical functions, about 100 pages, will be published separately and ready for distribution this fall. It is also announced that Parts III and IV will be published in a single book of about 284 pages. Part IV contains the seven place logarithms of numbers from 1 to 100,000.

NOMINATING COMMITTEE

The President of the Association has appointed the following members as a committee to nominate officers for the year 1925:

Leonard P. Ayres, Chairman
Wesley C. Mitchell
Warren M. Persons

The officers will be elected at the annual meeting at Chicago in December. This notice is given to meet the constitutional requirement that the names of our nominating committee shall be announced in the JOURNAL three months before the annual business meeting.

R. E. CHADDOCK,
Acting Secretary.

ANNOUNCING THE 1924 ANNUAL MEETING

The Eighty-sixth Annual Meeting of the American Statistical Association will be held at Chicago on December 29–31, 1924.

The experiment tried out last year of having the several sessions of the Annual Meeting deal with various phases of one general subject proved so successful that it has been decided to repeat the plan at the coming Annual Meeting of the Association. The President, after consultation with other members of the Association, has selected the topic "Population," and has developed a program on that theme. The subject is not only of great importance both theoretically and practically, but is one in which the statistician is in a strategic position to make a vital contribution to current political and economic thought. Moreover, it is a subject which bears directly upon the research activities of many of our members.

The subject of population is also arousing interest in allied scientific fields. For this reason it has been possible to arrange for joint sessions with the American Economic Association, and the American Association for Labor Legislation. With the economists we will discuss the connection between population and natural resources; with the labor legislation group the subject of population and the labor supply. In addition, we will develop independently three sessions, the first dealing with a statement of the problem and its various ramifications; the third dealing with the racial constitution of the population and the problem of immigration; and a final meeting in which we will discuss the outlook for the future, and the practical steps which should be taken in developing a rational public health, economic and social program.

As the subject is such a broad one and the time in which to cover it is so limited, the individual sessions are necessarily rather long and in some cases contain more papers than is altogether desirable. But it was necessary to arrange the program thus, in order to cover all angles of the subject and to round out every vital phase of the question. It is, therefore, impossible to appoint set persons to discuss each paper. But since many of our members are actively engaged in research work in this field, it should be possible to have a lively, even if it appear like an impromptu, discussion from the floor. There will be a reasonable length of time set aside expressly for such discussion and it is hoped that many of our members will show their interest by adding their contributions to the set papers which are scheduled.

Though the general outline of the program is fairly well settled, there is still time to make a few additions and changes and to this end suggestions and recommendations are earnestly requested. Information concerning original investigations now in process and the names of contributors who are doing valuable work on this subject

should be sent to the President. Above all it is hoped that many members will attend and take an active part in the discussion. The success of the Conference, as a whole, will depend upon the enthusiasm and whole-hearted support of each member.

LOUIS I. DUBLIN,
President.

New York City—September 10, 1924.

PRELIMINARY PROGRAM OF THE ANNUAL MEETING OF THE AMERICAN STATISTICAL ASSOCIATION, CHICAGO, ILLINOIS, DECEMBER, 1924

THE POPULATION PROBLEM IN AMERICA TODAY

President's Address—Monday Evening, Dec. 29

The statistician and the population problem

MONDAY MORNING, DEC. 29

I—The Statement of the Problem

1. The rate of growth of the American population
   Prof. E. B. Reuter, University of Iowa
2. The urbanization of the American population
   Prof. Hornell Hart, University of Iowa
3. The sources of natural increase of population
   Prof. Warren S. Thompson, Miami University
4. The limit of American population
   Prof. Howard Woolston, University of Washington

MONDAY AFTERNOON, DEC. 29

II—Population and Natural Resources (with the American Economic Association)

1. The economist's view of the population problem
   Prof. A. B. Wolfe, Ohio State University
2. Agriculture and population
   Prof. Don D. Lescohier, University of Wisconsin
3. The economics of the American farm
   Prof. David Friday, New School for Social Research
4. Increase in efficiency of farm production
   Prof. Alonzo E. Taylor, Stanford University
5. Mineral and other resources for future populations
   Dr. Frederick Gale Tryon, U. S. Geological Survey

TUESDAY MORNING, DEC. 30

III—Population and Immigration

1. The racial composition of the American population
   Prof. Henry Pratt Fairchild, New York University
2. The effect of infusion of alien racial stocks upon the development of an American type
   Dr. A. Hrdlicka, Smithsonian Institution
3. The contribution of alien heritage to national cultural development
   Dr. Alexander Goldenweiser, New School for Social Research
4. A rational American immigration policy; its basis and goal
   W. W. Husband, Commissioner General of Immigration
5. The immigration program of Canada
   R. H. Coats, Dominion Statistician

TUESDAY AFTERNOON, DEC. 30

IV—Population and Labor Supply (with the American Association for Labor Legislation)

1. The training and direction of our labor supply
   George Soule, Labor Bureau, Inc.
2. The wider occupational opportunities of women and the population problem
   Prof. Susan Kingsbury, Bryn Mawr College
3. Influence of child labor restrictions on population
   Miss Grace Abbott, U. S. Children's Bureau
4. The management of unemployment and its bearing in the population problem
   Bryce M. Stewart, Unemployment Insurance Office, Chicago

WEDNESDAY MORNING, DEC. 31

V—The Outlook for the Future

1. The family wage and the future population
   Prof. Paul H. Douglas, University of Chicago
2. The effect of the health movement on the population of the future
   Prof. S. J. Holmes, University of California
3. The essentials of an optimum social organization in the future population
   Speaker to be announced
4. Methods of approach to the population problem of the future
   Miss Mary Van Kleek, Russell Sage Foundation

MEMBERS ELECTED SINCE MAY

Achinstein, Asher, graduate student at Columbia University.
Ashbaugh, Charles C., Michigan Bell Telephone Co., Detroit, Mich.
Blumenthal, Frances, Public School No. 4, New York City.
Falkner, Helen D., Wheaton College, Norton, Mass.
Fall, Frank A., Ronald Press Co., New York City.
Fallon, William P., Western Electric Co., New York City.
Freudenthal, Elisbeth E., student at Harvard University.
Fukumoto, F., Tokyo Sericultural College, Tokyo, Japan.
Judson, Everett, American Radiator Co., Buffalo, N. Y.
Knapp, Arthur, United Gas Improvement Co., Philadelphia, Pa.
Leighton, J. Herbert, Industrial Dept., Federal Reserve Bank, Boston, Mass.
Loomis, Earl H., Western Electric Co., New York City.
McGaughy, James R., Teachers College, Columbia University, New York City.
Naughton, Joseph M., Eli Lilly & Co., Indianapolis, Ind.
Obenauer, Marie L., Industrial Survey and Research Service, Washington, D. C.
Rhoads, Willard R., Sanitary District of Chicago, Chicago, Ill.
Rider, Paul R., Washington University, St. Louis, Mo.
Saposnekow, Jacon, West Virginia University, Morgantown, W. Va.
Seeman, Albert L., Morningside College, Sioux City, Iowa.
Stafford, Thomas A., Methodist Episcopal Church, Chicago, Ill.
Suzuki, Tokitaro, student at University of Wisconsin.
Turno, Herbert C., Retail Chain Store Organization, New York City.
Tuthill, Stephen S., American Zinc Institute, New York City.
Varney, Mildred C., Bureau of Business Research, Harvard University.
Watrous, George D., U. S. Tariff Commission, Washington, D. C.
Weaver, Henry G., General Motors Corporation, Detroit, Mich.

REVIEWS

Wages in the Coal Industry, by John Wilkinson Rowe. London: P. S. King and Son. 1923. viii, 174 pp.

In the first chapters of *Wages in the Coal Industry*, and in fourteen notes in the appendix, Mr. Rowe pictures the varying methods of working and the innumerable factors of necessity and custom that make the methods of payment for coal miners in Great Britain vary from district to district, from mine to mine and even from one working place to another. If the book had no other object, an analysis of the complex factors upon which the wage structure is based would have been a contribution. This, however, is a by-product of the main discussion. The major problem deals with the effect upon earnings of the adjustment of wage rates by percentage increases or decreases on a standard rate. With the men who actually mine the coal paid on a piece basis, the standard can be nothing else than a piecework price list for hewers, and the actual day rates for day men at a given date. In adjustment, Northumberland, Durham, South Wales and Scotland each have separate percentage variations by contrast with the Federated Area for which there is normally but one percentage. In addition there is continuous individual and local bargaining to take care of changes in mining methods, bad working places, etc.

Over a long period, do the minute changes counterbalance, leaving the basic rates about the same, or do basic rates actually undergo a change though nominally the same? To answer this query, Mr. Rowe had the difficult problem of adjusting the classification of mine occupations in the Board of Trade Census of 1886 to that of the figures collected by Mr. Finlay Gibson in his schedules for the Mining Association of Great Britain in 1914. The later figures cover 80 per cent of the underground workers. The early figures represented 16 per cent of the workers in the industry—a meagre enough sample considering the wide differences in export and domestic markets, and in seam conditions, to say nothing of variation in job terminology and content with the changes in working during the period.

Having determined upon wage statistics per shift for two widely separated dates, the author calculates the amount of the percentage increases upon the adjusted 1888 figure in order to determine whether the computed figure approximates the actual wages of 1914. In this way, he finds that for tonnage men Lancashire and Lanarkshire, and a few of the smaller fields, "show no considerable difference between the calculated and actual wages in 1914." Despite great variations in the level of wages at different pits, the wages of piecework coal getters moved with "the recognized percentage alterations."

Combining the figures for all districts and taking tonnage men and firemen as representative of skilled groups, putters (car pushers) of semi-skilled, and laborers of unskilled, he finds that "skilled time workers' rates lagged considerably behind the advance in piecework earnings." The semi-skilled lagged also

though not to the same extent, while the unskilled gained about the same as the skilled; "The pieceworkers had secured their increase as a result of a variety of causes—successful bargaining over new price lists, more efficient tools and methods of working, and perhaps too in many districts as the result of shortened hours of labor." The analysis of the pre-war period shows, "that there is no ground for the assumption that over a long period, wages are regulated exclusively by the formal percentage variations; that, on the contrary, there is very often a wide discrepancy; that there has been no case of a decrease in basic rates, at any rate over the period as a whole; that there has been little or no uniformity in the amount of the extra increase secured, as between different districts, or between different grades, though it may perhaps be said that the skilled and semi-skilled day-rate workers have not benefited so much as the skilled pieceworkers." The analysis might be greatly complicated by the conditions of export trade in the years selected, seasonal fluctuation in the house coal trade, variations in wages paid in kind, and in mine practices concerning "occupational expenses," all of which the author recognizes even when the data do not allow accurate evaluation.

The movement of wages in the war period was further complicated by the system of flat-rate advances. Between 1914 and September, 1917, "hewers" received an increase greatly in excess of the nominal advance on the standards; unskilled men got little increase—in no case, except in North Wales, as much extra as the hewers. By November, 1918, the increases of unskilled workers "caught up with the skilled pieceworkers in almost all districts and passed them in the more important." In this period "the semi-skilled and unskilled had secured a slightly greater percentage advance than the skilled workers. This gap was very much widened by the further flat-rate advance in January, 1919, reduced slightly by the percentage increase of March, 1920, but was again increased greatly by the flat-rate advance of November, 1920."

Beset with difficulties as is the subject this careful treatment should warn anyone against computing earnings by an easy way from statistics of rates. Even the highly organized coal industry of Great Britain is not a unit within which a single workable agreement has prevailed. Economic factors have made for different wage bases within the different fields and for a wide range in earnings for workers in the same grade within a field.

The success of the analysis of methods of remuneration compensates the reader for piecing together the information scattered throughout the various chapters. The long-wall methods of mining and the small percentage of coal undercut by machine make impossible a comparison of methods of payment in bituminous mining in America. In fact, more likeness would be found in the structure of wages in the anthracite industry. A study of the steps by which miners' helpers in Great Britain came to be paid so largely on a day basis should yield some suggestions for the adjustment of contract miners' laborers' wages in the anthracite industry in Pennsylvania where the laborer is still largely paid a piece rate dependent upon the earnings of the contract miners.

ANNE BEZANSON

University of Pennsylvania

Railway Rates and Cost of Service, by Owen Ely. Boston and New York: Houghton Mifflin Company. 1924. 148 pp.

This volume was awarded one of the Hart, Schaffner and Marx prizes for the year 1922. It is a stimulating essay on the theory of railroad rates and its publication will help to clarify current thought on the subject.

The range of topics considered is wide for so small a book, and consequently the treatment of each topic is sketchy. The first 39 pages are devoted to a review of the rate policies of the Interstate Commerce Commission. After a very incomplete consideration of cases making reference to the cost basis as applied to individual rates, two chapters are given to the general rate cases, including *Increased Rates*, 1920, but not including *Reduced Rates*, 1922. The author recognizes the two widely different aspects of cost of service as related to railway rates: (1) cost viewed as the measure of the proper total income of a railroad, and (2) cost viewed as a measure of differences in rates as between specific services. The first aspect involves the problems of fair valuation and of judging of the economy and efficiency of operation; the second aspect involves such problems as the practicability of cost apportionments and of the economic effect of cost rates on commerce and industry.

The next 40 pages are devoted to the theory of railroad rates. The author draws a distinction between the study of the individual rate or tariff and the study of "group" rates, the latter dealing with "the set of principles by which the individual road makes its rates." One may question the usefulness of such a division of the field of rate theory. The author finds little justification in economic theory for value of service as a basis for rate making. He also attacks the joint cost theory of Professor Taussig and rightly points out that there is not a good analogy between the production of various kinds of railroad services and the production of gas and coke viewed as joint products. He says regarding the alleged fixity of joint costs:

> To assume that a large part of total cost is fixed, and will not vary with the amount of traffic, obviously fails to take into account railroad dynamics. As any railroad grows in size, its interest charges and its maintenance costs must increase. A small local road may operate without substantial change for a period of years, but the large systems typical of present day railroading must continue to increase their facilities year in and year out, the process of growth being fairly constant, though dependent to some extent upon industrial conditions and money rates.

He further points out that the reduction in railroad costs in the United States over a period of 50 years was not mainly due to the "fixed cost" principle, but to the principle of increasing returns, which applies to direct as well as to joint cost. This principle, he says, has nearly ceased to operate for some of the largest roads. He does not reach any precise theory of his own. He is led to build up a composite theory by a series of statements about rate making principles. In general he regards a cost system as an ideal basis for rates in a well-balanced industrial system, but he does not advocate an immediate and drastic modification of our freight rate system. In fact, he goes so far as to advocate the use of rate adjustments in serving national policies deemed essential to greater eco-

nomic and social welfare, as, for example, checking the growth of cities and the stimulation of agriculture.

One chapter is devoted to a statistical study of the law of increasing returns in transportation. It is shown that with the increasing traffic density from 1890 to 1920, there has been a marked fall in the average revenue per ton-mile when this is expressed in terms of purchasing power. It may be remarked that if we also express the railroad net operating income in terms of purchasing power, it will be apparent that the reduction in ton-mile revenue reflects not merely lowered operating costs but also in recent years a progressive reduction in the "real" earnings for the capitalist. It is also shown by a series of charts that rates in a given year are likely to be lower, within limits, on roads of greater density than on those of lighter density of traffic, and lower for the longer than for the shorter hauls.

Incidentally it may be noted that on page 89 a definition is given for "operating income" which is in fact that for "*net* railway operating income" as these terms are officially used.

The review of attempts to develop a specific cost formula for testing discrimination shows that much still remains to be done before we can base rates generally on cost even if we wish to do so. Perhaps the chief reason why cost of service accounting is not further developed is the fact that there is as yet no general acceptance of cost as the controlling factor in deciding most rate situations. The author refers to the fact that there are not available statistics showing the ton-miles and average ton-mile revenue for all commodity groups for which tonnage figures are now reported. It may be well to note that such information is now regularly compiled by certain roads, and plans for obtaining it generally have been worked out, but doubt as to whether the resulting statistics would be worth the added clerical cost, which would amount to over $62,000 a month, has so far prevented their adoption.[1]

M. O. Lorenz

Washington, D. C.

The Calculus of Observations, by E. T. Whittaker and G. Robinson Blackie and Son. London: 1924. vii, 389 pp.

Statistical methods or mathematical statistics are almost everywhere in America considered as something very modern and new. The truth is of course the opposite. Instead of being an infant in swaddling clothes mathematical statistics is a conservative method of investigation and of aristocratic lineage and can count among its ancestors several of the most distinguished names in the history of mathematical sciences, such as Pascal, Fermat, Bernoulli, de-Moivre, Newton, Lagrange, Cauchy, and above all the two great Frenchmen, Laplace and Poisson. In Bernoulli's famous *Ars Conjectandi*, which was issued long before the American revolution, practically all the elementary theorems of the so-called "modern" statistical methods are stated in a more rigorous and exact manner than in most of the recent crop of statistical text books. The

[1] See *38th Report of the Railway Accounting Officers Association* 1922, pp. 76–81 and pp. 132–135.

same holds true for deMoivre's remarkable "*Doctrine of Chances.*" Laplace's monumental work on the theory of probabilities still remains today, more than one hundred years after its appearance, by far the best text book on statistical method.

The above remarks will be sufficient to prove that a well-developed statistical technique has been available for nearly 200 years, long before we had extensive collections of crude statistical data. Statistical methods and mathematical statistics are in the final instance nothing more than ordinary theory of probability as taught by Laplace and Poisson. Unfortunately, however, for nearly a whole century the great mass of information contained in Laplace's work remained practically a closed book to English, American, and German statisticians, and it remained to the Scandinavian scholars, Gram, Thiele and Charlier to take up the splendid heritage left by the great Frenchman. (An exception is offered in the *History of Probability* by Todhunter.)

Perhaps one of the chief reasons for this neglect by the English speaking nations is found in the fact that everything French was looked upon with suspicion in England after the Napoleonic wars; and Laplace had dedicated his *Probabilities* to Napoleon.

The fact that the real worth of probabilities is as yet not recognized here in America is made manifest by the recent publication of the National Council of Research of its *Handbook of Mathematical Statistics*, edited by Professor Rietz. Fully two thirds of this book is nothing but a collection of probability formulas, but the theory of probability is treated in less than two pages, which contain some colossal blunders and gross misconceptions about the fundamental ideas of this science, such as for instance the awful blurt on page 16 in respect to the addition theorem.

I have purposely mentioned this *Handbook of Mathematical Statistics* in order to compare it with the present work by Professor Whittaker, because both books cover practically the same ground, but with this important difference that whereas the American handbook is nothing but a mere collection of diverse formulas without any adequate proofs, Professor Whittaker's treatise offers a detailed presentation with rigorous proofs of the various theorems. This is by no means a mean accomplishment. The American handbook gives about 200 pages of formulas *without* proofs. Whittaker supplies in 380 pages a greater number of formulas and gives rigorous demonstration of these formulas.

Professor Whittaker, who is one of the leading British mathematicians, has called his book *Calculus of Observations*, following in this respect in the footsteps of the Scandinavian astronomer and actuary, Thiele, who preferred to speak of the theory of observations instead of the theory of statistics, mathematical statistical or statistical methods. Personally I agree with both Thiele and Whittaker in their nomenclature.

Moreover, Professor Whittaker, like his Scandinavian colleagues, follows closely in the footsteps of the classical masters and in his *Calculus of Observations* lets the theory of probability be one of the fundamental factors.

The first four chapters which previously have been issued as a separate book under the title *A Short Course in Interpolation* gives undoubtedly the best text

which we possess in English on this very important subject. It is almost uncanny to note how this erudite Scotchman has been able to include practically all the important theorems on 52 pages together with a number of numerical examples. He has studied the old masters, Newton, Gregory, Euler, Briggs, Lagrange, Stirling and Bessel as well as their more modern commentators and gives a very lucid and compact representation of their work. In this connection it is especially gratifying that the author has treated the divided difference formula originally introduced by Newton and later on made the basis of the very general interpolation theory of the Danish scholars Opperman and Thiele whose work evidently has influenced Professor Whittaker's treatment. The idea that Newton's, Bessel's and Stirling's formula apply to equidistant intervals only is erroneous. By introducing the concept of divided differences we are able to operate with arbitrary intervals in even a more effective manner than the one supplied in the formula of Lagrange, which by many has been regarded as the only one applying to arbitrary intervals. About the only worthwhile interpolation which is missing in these four chapters is the powerful method of interpolation by means of orthogonal functions as introduced by the Danish actuary, Gram.

The 25 pages of the fifth chapter gives an elementary account of the important subject of determinants which no doubt will play a great rôle in the calculus of observations in the future. Chapter VI, dealing with numerical solutions of various equations is perhaps the weakest, but also the least important in the book. The seventh chapter is a continuation of the interpolation theory and deals with mechanical quadrature and numerical integration.

We now come to those chapters which I regard as the strongest in the book and which treat of normal frequency distributions and the method of least squares. Whittaker takes here unqualifiedly the standpoint of Laplace and Poisson and their Scandinavian followers, Gram, Thiele and Charlier, although he also gives much space to the far-reaching researches of Gauss. The main point to be emphasized is, however, that he bases his whole presentation on the probability theory of Laplace. No mention is made in any place of the Pearsonian system of frequency curves. A very pretty demonstration of the generalized frequency curves is given by means of Thiele's semi-invariants. The author shows in his footnotes that he is familiar with the original pioneer work by Gram in Danish in 1879 and the following research by Thiele and Charlier. Gram is specifically mentioned as being the first to find the general serial expansion of arbitrary frequency distributions. Gram's work in this respect is about twenty years ahead of the more recent investigations by Charlier, whose chief distinction is to have introduced the B or what I prefer to call the Poisson-Charlier series in distinction to the Laplacean-Gram or Gram-Charlier series.

Professor Whittaker does not follow up the generalized frequency curve. He limits—as he states himself—his investigations to the normal curve. But this treatment is rigorous and exact and at the same time brief. To dispose thoroughly of such an important subject in 35 pages is an accomplishment which I challenge any one of our dilettante writers on statistical subjects to duplicate. Similar words of praise apply to the chapter of least squares which occupies but

50 pages but is in my opinion superior to the massive volumes on the same subject by several American and German writers. The only criticism I have is that the author in the solution of normal equations seems to rely almost exclusively on determinants. I will admit that the theory is elegant, but its application in practice sometimes meets with obstacles, and I should like to have seen a somewhat more detailed statement of Gauss' method of solving simultaneous equations than that which is found in the book.

Chapter X is a chapter I advise our friends the economic statisticians to study very carefully. It covers only 25 pages. But I dare say it will require much study on the part of at least 99 per cent of our *Gelehrten* among the professors in economy and statistics to master this chapter on Practical Fourier Analysis. I have myself already spent a full month thereon and have only finished 18 pages. It would be false modesty on my part to say that my friends among the economists and the majority of practical statisticians know more about this subject than myself, and when it will take me nearly two months to digest the highly concentrated information which Professor Whittaker has brought within the compass of this chapter, I am probably not far from the truth when I recommend the *Gelehrten* to take a year of sabbatical leave for a thorough intellectual rumination of Mr. Whittaker's nutritive diet. I feel sure that the result will be highly beneficial in so far as their future studies are concerned and will reduce their printer's bills by fully 50 per cent.

Similar remarks hold true for the chapter on correlation. Correlation has at present an almost magical effect. A hard-working horde of *Gelehrten* in various branches of learning are calculating coefficients of correlation between all phenomena their fancy happens to strike. The results are really love's labour lost. The Danish actuary, Jörgensen, in his great work on *Frequency Surfaces and Correlation*, has declared that the concept of correlation is as yet in a nebulous state and that it will take the diligent, meticulous and serious labors of many mathematicians to build even a firm basis for the theory. So far, the only sound basis is that supplied in the case of normally distributed variates, and even here there is much to be developed. A reading of Whittaker's chapter on the subject confirms this view by the Danish scholar. Let the mathematicians first weld an effective tool to attack the so-called correlated phenomena. Then it will be plenty of time for the statisticians to use it. At present they are only playing with a fascinating and somewhat hilarious intellectual gas, which is better left alone until the mathematicians have reduced it to a more solid substance.

The book is the result of a series of lectures given at the Mathematical Laboratory at the University of Edinburgh and is therefore designed to meet the application of mathematical analysis to concrete problems in practical life. The following quotation from the preface is of interest: "When the Edinburgh Laboratory was established in 1913, a trial was made, as far as possible, of every method which had been proposed for the solution of the problems under consideration, and many of these methods were graphical. During the ten years which have elapsed since then, the graphical methods have almost all been abandoned, as their inferiority has become evident, and at the present time the

work of the Laboratory is almost exclusively arithmetical. A rough sketch on squared paper is often useful, but (except in Descriptive Geometry) graphical work performed carefully with instruments on a drawing board is generally less rapid and less accurate than the arithmetical solution of the same problem." I beg to submit this erudite statement to those statistical authors who at present are so busily engaged in issuing bulky texts on graphical methods.

This review is probably already too long. Let me therefore in conclusion state that in my opinion Whittaker's treatise represents one of the few worthwhile presentations of what is called mathematical statistics. The chief reason perhaps being that his study is based upon the concepts of the classical writers on probability. Mathematical Statistics is but a synonym for the older word, Probability.

ARNE FISHER

New York City
June 29, 1924

Tenth Annual Report of the Federal Reserve Board Covering Operations for the Year 1923. Washington: Government Printing Office. 1924. 80 pp.

This tenth annual report of the Federal Reserve Board is a document of unusual interest to the economist and especially to the student of banking. Its interest is found predominantly in the fact that while the report in question furnishes all of the usual data concerning the operations, earnings and general transactions of the Federal Reserve System for the year 1923 so that it is statistically even more complete than usual, it also returns to an older custom which of late years has been rather more honored in the breach than in the observance. This is the practice of devoting a substantial amount of attention to broad, general questions of theory and of the money market. In past years, the Board's report has shown fluctuations of feeling between the ideal of a "business man's report," terse and to the point with little dreamed of in its philosophy except earnings, losses and "meeting the situation" in general, on the one hand, and the ideal of economic discussion on the other. There is a middle ground between the two extremes, in which the outlines of the general underlying conditions that obtain in the financial market at any time may be recognized, and a careful analysis of them may be given in broad terms informed, so far as possible, with a sound economic philosophy. This report for 1924 shows some disposition at least toward acceptation of the latter view, and is correspondingly to be commended by students of the subject of banking in its larger aspect.

That part of the Board's report which is thus of singular interest has to do with the general principles which have controlled the discount policy of the System and with the theory of the relation between that policy and general open market rates and conditions. It is particularly striking in that it represents one of the few occasions upon which the Board has been willing to state a theory of discount rates. Governors of the Board, governors of reserve banks and many others have from time to time presented their views on this subject either before congressional committees or elsewhere and have not hesitated to adopt very

conflicting points of view. Indeed there is hardly a theory of discount rates which has not been accepted and advocated by some responsible person representing the System. But the Board itself has been far more inclined to hesitate.

In the present report, an effort is first made to furnish an account of the discount policy during 1923. There were fewer "changes than in any other one year in the history of the System," and hence the policy of that year appears to have been negative, consisting chiefly in standing by and doing nothing unless action should be demanded. Sharp movements of prices are admitted to have taken place within that time, but the Board was of the opinion that they were not due to "unwarranted" changes in the amount of credit outstanding. From this discussion of recent events, the Board shifts to a consideration of the general question of relations between discount and market rates. It first calls attention to the obvious fact that the "margin between the Federal Reserve bank rate and member bank rates varies somewhat from district to district and . . . is narrowest in those districts where the financial centers are located." It then rejects the principle adopted in England of having the bank rate "lead" the money market, reaching the conclusion that "English banking practice does not . . . establish the inference that Federal reserve discount rates, in order to be effective, must be penalty rates—that is, be higher than the rates charged by member banks on customer loans." This makes it necessary for the Board to substitute a theory of discount rates. But it finds that "the experience of the Federal Reserve banks under normal conditions of operation has as yet been too brief to make it possible to speak comprehensively and definitely concerning the proper relation of Federal Reserve bank discount rates and member bank customer rates." However, a few generalities are possible. One is that occasionally Federal Reserve rates may be an important and at times a leading influence in money centers. Another is that changes in prices do not afford a good guide to policy in the changing of discount rates. Still another is that statistical indices of production and the like come too late to serve as a guide in this particular. So, the Board is finally driven to the conclusion that at the present time there is nothing except "discretion" which affords a satisfactory basis for the fixation of discount rates. Just what discretion is, and how it is attained, is left as uncertain as it has always been in the history of human thought.

One favorable indication in the report is that the Federal Reserve Board has apparently aroused itself to the necessity of developing the open market operations of reserve banks. It descants a good deal upon the importance of these operations, and remarks that "the experience of the past year in open market and discount operations . . . is significant." In fact, the Board, or those who wrote this report, have apparently discovered the importance of the open market function almost for the first time, and put forward the well-worn conclusion that it affords a very useful way of adjusting the supply of credit to the condition of the market, instead of waiting for the presentation of member bank paper for discount. "There can be no doubt that the Federal Reserve Act looked forward to the development in the United States in the course of time, of an open market of considerable extent . . . ," says the Board sapiently.

No! there can be no doubt on that subject. In fact, the Act is practically mandatory on the point, as a re-reading of its terms would show, and the moment when the development was expected to begin was not "in course of time," but at once. The System has waited ten years to undertake any such dealings on an extensive scale and it may be doubted how far the managers mean to go at the present time. Still it is a hopeful indication that the importance of the subject is at length realized.

After some years, an annual report of the Federal Reserve Board thus recurs to some of the really fundamental problems of the System. It has little to say about branch bank buildings at Podunk Hollow or the vagaries of Senator X and his animadversions upon the members of the Board. This at all events is decided progress and should be recognized as such.

H. Parker Willis

Miner's Wages and the Cost of Coal, by Isador Lubin. New York: McGraw-Hill Book Company. 1924. 316 pp.

This is the kind of book that clarifies issues in controversy and makes progress toward solutions. It is impressively different from the labor histories with which we have been made familiar, the partisan narratives of the battlefield written on the assumption that moral issues and such issues alone, are at stake. In this admirable study the main stress is laid on basic faults of organization in the coal industry, on shortsightedness and confused economic thinking on the part of both operators and union officials, and to both contending sides much food for thought is given.

Half of the book is devoted to the machinery for making wage rates and to the theory, more or less consciously evolved, by which rate making proceeds. The rise of the miners' union to great power is sketched and contrasted with the disunion of the operators. The rise of the Interstate Joint Conference is described and its relation to the conferences of lesser jurisdiction is discussed. What was at first manifested incidentally as a device for securing the widest possible adherence of operators to the negotiations of the Joint Conference became in time a fixed principle of predominant importance: the principle of "competitive equality," a principle which sought to make the employers equal before the workers, but which in practice gave rise to and perpetuated amazing incongruities which in the long run unsettled rather than stabilized the industry. The chapters on pick mining rates and on machine rates show the length to which these incongruities were carried. In various ways, but not uniformly, the rates made for the Central Competitive Field have influenced those made in the outlying fields. What gains are achieved by one side or the other, however unsuitable they may tend to become in time, are not changed until an access of bargaining strength comes to the other side.

The inherent difficulties in the theory of competitive equality are portentous, but the application of the theory bristles still further with inconsistencies. The chapters of this book in which the various differentials are discussed are an acute analysis of their topics and deserve a most careful reading.

The book purports to show the relation of miner's wages to the cost of coal. Pertinently the author points out that various other considerations go farther in explaining fluctuations in price than do changes in wages. Year in and year out, however, the differentials for machine, for freight, and for thin seams cause an uneconomic exploitation of our coal resources which is inevitably reflected in coal prices.

Mr. Lubin succinctly summarizes his thought in this connection (page 289):

> (1) In so far as the wage system provides relatively lower wage rates for disadvantageously located and thin-vein properties, it makes possible the operation of mines which have no economic justification. (2) The wage system to a degree, then, is responsible for the overexpansion that today characterizes the bituminous coal industry. (3) This overexpansion has brought with it irregularity of operation, excessive costs, and waste of our coal resources. It accounts in large part for the annual loss of over 30 per cent of the miner's possibile working time; it makes necessary the maintenance of some 200,000 surplus miners; it adds to the consumers' coal bill some $100,000,000 of overhead charges for periods of shutdown; and it leads to the wasteful exploitation of our coal fields. (4) These evils can be eliminated in some degree by a change in the wages system. The decision as to whether or not such a change will be made lies immediately with the mine owners and the union.

Only by elimination of the differentials can a remedy come, and the price of such an elimination would be high: the abandonment of millions of investment, the displacement of several hundred thousand men, the abandonment of hundreds of mining towns and the uprooting of thousands of families.

It must not be supposed that the reader who follows Mr. Lubin's logic will inevitably understand how to set the industry on its feet. There is the organization dispute to be settled. That being a study by itself, it is (except for the comment in a footnote on page 220) excluded from this book. There is the practical problem of the elimination of the differentials—some methods must be more merciful than others—but that the Institute of Economics will doubtless discuss in another volume. There remains also the problem of differences in wage levels which do not depend on competitive inequalities but on regional differences of various kinds—on such differences, for example, as were recognized by the British Coal Mines Minimum Wage Act of 1912.

ROBERT F. FOERSTER

Princeton University

Principles of Vital Statistics, by I. S. Falk, Ph. D. Philadelphia: W. B. Saunders Company. 1923. 258 pp.

A text, such as Dr. Falk has supplied, undoubtedly was needed. The more comprehensive treatises by Newsholme, Whipple, and Pearl have important places on the desks of health officers and vital statisticians, but it has been found difficult to utilize any of these three excellent books for short courses which are being given to an increasing extent to workers in the field of public health who are not *primarily* concerned with the collection or analysis of vital statistics. What was wanted was a very simple yet sufficiently comprehensive introduction to the subject, rather than to the methodology, of vital statistics which could be

covered by students of this kind in a comparatively short time. Professor Winslow, in his foreword, refers particularly to courses for public health nurses as an illustration; a number of other kinds of workers in the public health field might be mentioned, such as inspectors in municipal or county health departments, who are in need of a presentation such as Dr. Falk gives.

There is one feature of Dr. Falk's text which, in the reviewer's opinion, should be commended especially. This is the emphasis which he gives to morbidity statistics. Approximately one fifth of his book is devoted to the subject of morbidity. He points out, with proper emphasis, the importance of securing "accurate data on the extent of sickness in ordinary populations; in special types of population; the severity of sickness; the available organizations to care for the sick; the causes; the forms of sickness; its incidence by sex, age, color, nationality, and race." It might have been well to acquaint the group of persons for whom this book is primarily intended with the fundamental differences between mortality data and morbidity data in the analysis of public health problems, since the character of the two kinds of data is not always fully appreciated even by the more experienced workers in the field of vital statistics. Another important point in the chapters on morbidity might have been emphasized to greater degree, namely, actual practice in the notification of disease, and the availability of morbidity data from this source. There have been in the last few years some very important developments in this field which have gone considerably in advance of the so-called model notification law and standard notification blank, and in certain states and cities some very valuable contributions have been made which should be better known to those who expect to work in the field of public health.

In attempting to compress an adequate description of the data of vital statistics into a volume of 250 pages, however, the question of what to put in or what to leave out and of how much space to give to each topic must have been a very difficult one. It is very easy to criticize any attempt of this kind, and doubtless in the use of the book a teacher will be disposed to criticize it because some topic was not given sufficient space. There is no reason why such a teacher cannot supplement the book in the direction that he wishes. The text itself is an admirable presentation for its particular purpose.

EDGAR SYDENSTRICKER

Washington, D. C.

Commerce Yearbook, 1922, United States Department of Commerce. Washington: Government Printing Office. 1923. 692 pp.

This first issue of a series of annual volumes inaugurates a much needed addition to periodic literature on current economic phenomena in the United States. While there are many volumes containing quantitative data descriptive of the immediate past, terse yet careful and comprehensive discussions of those statistics are rare and widely dispersed. The authors of the new Yearbook have used the numerical indices of our economic existence as the basis of critical analysis of our national economic position, and have brought between two covers careful discussion of nearly all angles of the present day business situation. The field covered includes analysis of conditions which existed during 1922 and the early

months of 1923 in individual industries in the United States, in construction, agriculture, prices and volume of trade, transportation and communication, financing and banking, and foreign trade. An extensive section is devoted to international trade during that period and to economic reviews of the producing areas of the world. An elaborate and well executed synopsis of the volume is presented in the first twenty-six pages. This and the summaries in heavy type at the beginning of most of the sections on specific topics facilitate brief examination of the whole commercial field. The reader is enabled to scan quickly the parts in which his interest is general and to concentrate his reading on the special topic which most concerns him.

There is no attempt at mathematical treatment of the data. The mathematical statistician engaged in analysis of the time series has no greater difficulty to contend with than the urgent demand of business that, from brief series, he make accurate predictions of short time fluctuations. This from the very nature of his method he is unable to do, for numerous forces, in no way routine in their nature, play so large a part in determining the immediate future that mathematical treatment of them is extremely difficult. If short time forecasts made by mathematical economists from brief series are compared with deductions made from general experience and broad information by skilled empirical generalizers, it is found that as yet the latter method is in many instances much superior to the former.

Presenting in brief form carefully analyzed material of wide scope, together with deductions from that material, the Yearbook should be of real value to the clientele for which it has been prepared—in the words of the letter of transmissal, "American business men, bankers, economists, trade association executives and the trade press." For those engaged in statistical research in the fields of business and economics, the Yearbook is of little use, since except for illustrative tables it contains none of the wealth of tabulated material which makes the two sister series, the *Statistical Abstract of the United States* and the *Monthly Summary of Current Business* of such real worth. For the practical man in the business field this is a real advantage, for the mind is freed from consideration of lengthy tables containing data for many years past. The facts of the year and, for the sake of comparison, a few other years are brought directly and in simple form before the reader, and the analysis proceeds from them. While the value of careful analysis of lengthy time series is unquestioned, for the average man in practical fields much is to be gained from the less technical method of treatment. Simple charts add to the clarity of the work.

If criticism of the method of treatment is to be offered, it must be levelled at what appear to be certain rather dogmatic statements. It is evident that these arise from the wide field covered by the volume and the necessity for brevity. They are particularly noticeable in the general review that opens the volume. An illustration is found on page 6, "The most significant measure of the economic situation of a country is the volume of production." Again, on page 13, "Turning now to the machinery, locomotive, and car-building industries, we find a similar increased activity, *which was the result of reviving confidence.*" These can be little more than expressions of opinion.

The several sections evidently are the products of various writers, and it is to be expected that there will be a certain lack of uniformity in method of treatment and worth. Nearly all have been ably handled. One exception is the part dealing with the chemical industry which is very briefly and superficially discussed. Also certain fields have been entirely neglected, the most important of these being marketing.

The usefulness of the volume would be immeasurably heightened by a more carefully compiled and more elaborate index. It is doubtful if the clintele of the series will be nearly as much interested in an economic review of Iraq as in the production of glass or of asbestos in the United States, yet the former is indexed and the latter are not. Nearly half of the page references in the index refer to foreign countries.

The new Yearbook is a laudable effort and will undoubtedly meet with cordial and well-deserved appreciation in business circles. And as the successive issues are prepared, minor faults, such as those just cited, will disappear. It is to be hoped that publication will be regular and prompt as the value varies inversely with the time elapsed since the period discussed.

FRANK A. ROSS

Columbia University

Mortgages on Homes. Census Monograph II, Department of Commerce, Bureau of the Census. Washington: Government Printing Office. 1923. 277 pp.

This volume presents, with explanatory text, certain new data obtained in an inquiry in regard to encumbrances on homes (other than farm homes). The data are compared with statistics previously published in order to throw light on home ownership and other related problems.

Although few will deny the significance of a high percentage of home ownership in the United States as a measure of the effectiveness of our social and economic organization, and as a factor in promoting cultural and economic progress, it may not be out of place to quote Professor Richard T. Ely's statement in the foreword: "The percentage of homes owned is always a function of the industrial system and the totality of economic conditions which prevail at any time." Secretary Hoover has stated, "The present large proportion of families that own their homes is both the foundation of a sound economic and social system and a guarantee that our society will continue to develop rationally as changing conditions demand."

The problem of financing the acquisition of homes is of vital concern in the situation, for the great majority of families must borrow in order to buy a home.

The new figures relate solely to the homes, other than farm homes, owned by the occupants, reported in the Census of Population for 1920 as having a mortgage or other encumbrance upon them. For this class of homes was obtained: (1) the value of the home, (2) the amount of mortgage debt at the date of taking the census, and (3) the annual rate of interest actually being paid.

These figures are presented in combination with others showing home ownership, number of families per dwelling, the rate of growth of population in different

states and cities, and comparable figures for other census years. The Census of 1890 is the only other one for which figures are available on the value of encumbered homes (not on farms), the amount of mortgage debt on them, and the rate of interest.

A summary table for the United States showing on one page the essential facts in regard to the tenure, value, interest rates, ratio of debt to value, annual interest charges, etc., would have been helpful. It would also be interesting to see tabulations for conditions on the 4,797,000 homes which are in rural districts but not on farms, and for those in incorporated places of from 2500 to 25,000 inhabitants.

The foreword by Professor Ely indicates with sufficient thoroughness the limitations of the material presented and the need for an intensive study of related facts in selected communities "in order to test the accuracy of the statistics and discover the possible relationship between these many factors which may well be the more important causes for the conditions as set forth by the facts discovered in this report." As he points out, the age of the debtors, their occupations, their incomes, the length of time it takes to pay off the mortgage, the extent of borrowing on the home to obtain capital for use in business or to purchase automobiles, the extent of foreclosures, etc., would all be instructive if we had the data. The analysis covers such points as the fact that although the rate of interest reported sometimes represents the true rate, in many cases there are probably bonuses or commissions, and the face value of the loan may also be inflated to cover the discount which must commonly be accepted for second mortgages.

The fact remains, nevertheless, that the volume serves a very practical purpose to those who wish to do something constructive in regard to developing better financing machinery for home owners. It helps define the problem, nationally and locally. The statistics for each city having 25,000 or more inhabitants show the number of mortgaged homes valued at less than $2,500, $2,500 to $5,000, etc., the mortgage debt on each class of home, the average annual interest rate charged on mortgaged homes, the number and amount of mortgages classified by interest rates, and the percentage of home ownership.

With a knowledge of local conditions in any town and in a few other cities with which it is desired to make a comparison, the basis of an effective appeal for constructive measures can be made. Whether or not a final solution to the problem involves action on a national scale, a program for future action must be evolved from experience in individual localities. The publication of this handy volume is timely, and a helpful contribution in every way.

JOHN M. GRIES

Dep't of Commerce, Wash., D. C.

The American Labor Yearbook, 1923-24. Edited by Solon De Leon and Nathan Fine. New York: Rand School of Social Science. 548 pp. $3.00.

The fifth volume of this well-known digest of contemporary labor history more than maintains the high standard set by the preceding issues. The material presented really falls into two major divisions—American labor movements and foreign—and approximately equal halves of the book are devoted to each. The

discussion of American events is fresh and concisely written. There is a succinct statistical summary of present economic conditions while the description of the tangled web of labor's efforts in the political field is comprehensive and illuminating. The only criticism that might be advanced as regards the American section is that more space might profitably be devoted to the activities of the trade unions on the economic field.

The material upon foreign labor events is even more useful. The activities of the trade unions and the labor parties in no less than fifty-three countries are described in such a well balanced and concise way as to give the reader a bird's-eye view of what the present situation is and what the forces at work are. One of the most striking sections of this part of the book is that dealing with the International Relations of Labor. One is impressed with the growing internationalism of labor organization. In addition to the Amsterdam and Moscow trade-union Internationals, there are more than thirty trade internationals in such occupations as bookbinding, the building trades, mining, and transport work. It must be a matter of regret to all who are internationally-minded that the American Federation of Labor and the National Women's Trade Union League have withdrawn from affiliation with the Amsterdam International.

In brief, this Yearbook is invaluable not only to trade-union officials and to journalists and speakers but to the scholar as well. The Rand School and Messrs. De Leon and Fine are to be congratulated upon so accurate and complete a work.

PAUL H. DOUGLAS

University of Chicago

RECENT LITERATURE

BOOK NOTE

The Fixing of Wages in Government Employment, by E. Colston Shepherd. London: Methuen & Co., Ltd. 1923. xx, 207 pp.

An American reader would move to amend the title by inserting words to indicate that the book deals almost exclusively with wage fixing in the British national government. It does not touch on the practice of British local governments and it devotes but a single chapter of eleven rather small pages to a "brief survey" of the wage system of foreign governments.

The nature of the treatment can perhaps be best indicated by quoting the short preface.

"The chief object of this study has been to collect and analyze the available material bearing on the Government practice in settling wages. This is done in the four chapters dealing with the War Office, Admiralty, Civil Service, and Post Office, which constitute the main substance of the work. In the introduction an attempt is made to indicate briefly the constitutional aspect of the problem and its relation to current economic theory on the subject of wages. The changes of the war period are summarized in Chapter V, though no attempt has been made to deal with these in such detailed fashion as the pre-war conditions are studied; and the main conclusions are added."

To American readers the most interesting chapter is the one on changes of the war period, because it touches on Whitley Councils and employee representation in a service which in the past has been dominated by an autocratic administrative system with the Treasury at its head. That subject, however, is more interesting to the political scientist and the economist than to the statistician, a statement that may well be extended to the book as a whole.

LEWIS MERRIAM

Institute for Government Research

STATISTICAL BOOKS

American labor year book, 1923–24. Rand School of Social Science. Labor Research Department. New York, 1924. 548 pp.

DOWNEY, E. H. *Workmen's compensation.* New York: Macmillan, 1924. xxv, 223 pp.

ELY, OWEN. *Railway rates and cost of service.* Boston: Houghton, 1924. x, 138 pp.

FRY, C. LUTHER. *Diagnosing the rural church. A study in method.* New York: Doran, 1924. xxvi, 234 pp.

LUBIN, ISADOR. *Miners' wages and the cost of coal: an inquiry into the wages system in the bituminous coal industry and its effects on coal costs and coal conservations.* (Institute of economics series.) New York: McGraw, Hill, 1924. xiv, 316 pp.

MELLON, ANDREW W. *Taxation; the people's business.* New York: Macmillan, 1924. 227 pp.

NATIONAL INDUSTRIAL CONFERENCE BOARD. *Wages, hours and employment in American manufacturing industries, July, 1914–January, 1924.* New York, 1924. vii, 104 pp.

PASVOLSKY, LEO AND MOULTON, H. G. *Russian debts and Russian reconstruction.* (Institute of economics series.) New York: McGraw-Hill, 1924. 247 pp.

POWELL, JAMES E. *Payment by results: introduction, organization, rate-fixing.* New York: Longmans, 1924. viii, 411 pp.

RIETZ, H. L. Editor. *Handbook of mathematical statistics.* Boston: Houghton, 1924. viii, 221 pp.

SCOVELL, CLINTON H. *Interest as a cost.* New York: Ronald, 1924. v, 254 pp.

STELLA, ANTONIO. *Some aspects of Italian immigration to the United States; statistical data and general considerations based chiefly upon the United States censuses and other official publications.* New York: Putnam, 1924. xxii, 124 pp.

WRIGHT, PHILIP G. *Sugar in relation to the tariff; with the aid of the council and staff of the Institute of economics.* New York: McGraw-Hill, 1924. xiii, 312 pp.

GOVERNMENT PUBLICATIONS

UNITED STATES.

——Bureau of Labor Statistics.

Employment in Selected Industries. April and March, 1924. 12 pp. each.

——Bureau of the Census.

Cotton Production and Distribution, Season of 1922–23. 103 pp.

——Railroad Labor Board.

Wage Series: Report No. 4. Average Daily Wage of Railroad Employees in Class 1 Carriers. 12 pp.

——Federal Reserve Bulletins:

May 1924. "Reparation payments and stabilization of German Currency." "Business Conditions in the U. S." "Report of Committees of experts to Reparation Commission."

June 1924. "Money Rates and the Demand for Credit." "Business Conditions in the U. S." "Resumption of gold payments by the Swedish Riksbank." "New Regulation on Check Collections."

——Public Health Reports.

No. 14. "Provisional Mortality and Birth Rates, First Nine Months, 1923. 80 pp.

No. 15. "Some tendencies Indicated by the New Life Tables." 52 pp.

No. 16. "Sickness Survey among 21,000 Automobile Workers." 68 pp.

No. 17. "Factors in the Mental Health of boys of foreign parentage." 60 pp.

No. 18. "The Notifiable Diseases in small cities, 1922."

No. 22. "Extent and Trend of Drug Addiction in the United States." "Summary of provisional birth and mortality figures, 1923." 80 pp.

No. 23. "The Public Health Problems of Chile. 65 pp.

No. 24. "A Study of Sickness in South Carolina Mill Communities." 73 pp.

State of Washington.

Department of Efficiency. Supplement No. 2 to Financial History.

State Current School Fund History, Period Statehood to 1923.

FOREIGN COUNTRIES.

UNION OF SOUTH AFRICA.

Office of Census and Statistics. Statistics of Production . . . for the Year 1921–1922. 73 pp.

New South Wales.

The Official Year Book Of New South Wales, 1923. Part I. Geographical History and Government, 44 pp. Part II. Education. 34 pp. Part III. Public Finances. Railways and Tramways. 62 pp.

Bureau Municipal de Statistique d'Amsterdam. No. 72.
Le Recensement de la Population. 39 pp.
Commercial Intelligence Department, India, in 1840.
Revies of the Trade of India, 1922–1923. 86 pp.
New Zealand. Census and Statistics Office, Wellington. Statistical Report on the External Migration of the Dominion of New Zealand for the Year 1923. 28 pp.
——Official Year Book 1924. 782 pp.
Western Australia. Statistical Register of Western Australia for 1922–23 and previous Years.
L'Office de Statistique de la Republique Tchecoslovaque. Prague 1923.
Commerce Exterieur de la Republique Tchecoslovaque en 1921. 244 pp.
Egypt, Ministry of Finance. Cairo: 1923.
Annual Statement of the Foreign Trade of Egypt during 1920–1922. 550 pp.

ARTICLES IN PERIODICALS

THE ANNALS.
May, 1924. Competency and Economy in Public Expenditures. Edited by Clyde L. King.
POLITICAL SCIENCE QUARTERLY.
June, 1924. H. S. Quigley: "Some Aspects of China's Constitutional Problem." H.S. Knapp: "Treaty No. 2 of the Washington Conference." A. G. Dewey: "On Methods in the Study of Politics, II." B. Stern: "The present Status of the Labor Movement in Germany."
THE JOURNA OF POLITICAL ECONOMY.
June, 1924. J. L. Laughlin: "The Quantity Theory of Money." W. R. Camp: "Agriculture and Price Stabilization I." E. Frankel: "The Eight-Hour Day in Germany." E. E. Witte: "Value of Injunctions in Labor Disputes." H. G. Brown: "Is a Tax on Site Values Never Shifted?"
MONTHLY LABOR REVIEW.
May, 1924. "What 86 Years Have Taught Us About Selected Labor." "Baltimore and Ohio Experiment in Industrial Coöperation." Index to Volume XV, July to December, 1922.
INTERNATIONAL LABOUR REVIEW.
May, 1924. F. Sitzler: "The Adaptation of Wages to the Depreciation of the Currency in Germany." E. Wigforss: "Industrial Democracy in Sweden."
March, 1924. P. Fauchille. "The Rights of Emmigration and Immigration." J. B. Condliffe: "Experiments in State Control in New Zealand." W. S. Parker: "The Problem of Seasonal Unemployment in the Building Industry."
INSTITUTE INTERNATIONAL DE STATISTIQUE.
Bulletin Mensuel de L'Office Permanent. March, April and May, 1924.
METRON. I–II,1924. M. Ptucha. "Die Sterblichkeit in Russland.
JOURNAL DE LA SOCIETE DE STATISTIQUE DE PARIS.
May 1924. Proces-Verbal de la Seance du 16 Avril, 1924. M. D. de Bernonville: "Les Indices du Mouvement General des Prix en France."
LE MUSEE SOCIAL.
April, 1924. A. Siegfried: "L'angleterre d'Aujord' hui."
June, 1924. E. Martin-Saint-Loon: "Le mouvement syndical europeen."
GIORNALE DEGLI ECONOMISTA.
March, 1924. G. Del Vecchi: "Stringere I Freni." E. Corbino: I Porti della Sicilia." A. Loria: "A Mia Difesa."

April, 1924. C. B. Turroni: "Studi sul Deprezzamento del Marco." Tedesco U. Ricci: "La Limitazione dei publici esercixi e l'alto costo dei viveri." M. Boldrini: Il "Trattato delle reliquie di giovanni." G. Jaia: "Per un Trattato di Geografia."

THE JOURNAL OF ACCOUNTANCY.

May, 1924. Prize Essay by S. Gundelfinger.

June, 1924. Prize Essay, part II. A. B. Barakat: "Linoleum Manufacturing Accounting."

JOURNAL OF THE INSTITUTE OF BANKERS.

May, 1924. J. Brunton: "Statutory Limitation of Actions." G. D. Chisholm: "Stabilization of Property Values."

JOURNAL OF THE INSTITUTE OF ACTUARIES.

March,1924. W. P. Elderton and H. J. P. Oakley: "Notes on the Interpretation of 'Select' Rates of Mortality." L. M. Laing: "Notes on the Industrial Assurance Act, 1923."

PUBLIC HEALTH.

May, 1924. C. R. Gibson: "The Prime Factor in the Decline in Infant Mortality."

AMERICAN ECONOMIC REVIEW.

June, 1924. W. Orton: "Social Philosophy of Mr. Bertrand Russell." C. E. R. Sherrington: "Results of British Railways Act of 1921." W. C. Mitchell: "Commons on the Legal Foundations of Capitalism." J. Bauer: "Recent Decisions on Valuation and Rate Making."

CASUALTY ACTUARIAL SOCIETY. PROCEEDINGS.

November, 1923. H. E. Ryan: "The Society and its Relation to Rate-Making Associations." J. D. Craig: "Allocation of Expenses." H. O. Van Tuyl: "A New Experience Exhibit for Casualty Insurance Companies." S. D. Pinney: "Miscellaneous Property Damage Insurance."

MISCELLANEOUS REPRINTS AND PAMPHLETS

SHIFTING OF OCCUPATIONS AMONG INDUSTRIAL INSURANCE POLICYHOLDERS, by Louis I. Dublin and Robert J. Vane. Reprinted from the Monthly Labor Review, April, 1924.

AN IMPROVED METHOD OF COMPUTING METEOROLOGICAL NORMS, by H. W. Clough. Reprinted from the Monthly Weather Review, Aug., 1923.

Statistical Bulletins of the Metropolitan Life Insurance Co.

TUBERCULOSIS DEATHRATE AMONG INDUSTRIAL POPULATIONS OF AMERICAN CITIES AND CANADIAN PROVINCES. April, 1924. Vol. 5, No. 4.

DISEASES WHICH KILL MORE WOMEN THAN MEN. May, 1924, No. 5.

BELL TELEPHONE SECURITIES, REFERENCE TABLES AND DESCRIPTIONS. American Telephone and Telegraph Company and Associated Companies. May 1, 1924.

New York Association for Improving the Condition of the Poor.

THE HEALTH OF A NEIGHBORHOOD, by John C. Gebhart.

TUBERCULOSIS, A FAMILY PROBLEM, by John C. Gebhart.

CRITERIA OF INVESTMENT, address to the Undergraduate students of engineering and economics, by George H. Davies. March 1, 1924.

EXPENSES, PROFITS AND LOSSES IN RETAIL MEAT STORES, by Horace Secrist. Series III No. 9. Northwestern University School of Commerce, Bureau of Business Research.

COMPARISON OF AN INDIVIDUAL CONCERN WITH THE HARVARD INDEX OF GENERAL BUSINESS. Harvard Economic Service, 1924.

HUMAN CAPITAL, by Charles H. Wickens, From Report of the Australian Association for the Advancement of Science, Vol. 16, pp. 536–554, 1923.

A BRIEF STUDY OF STATE DISTRIBUTION OF COLLEGE STUDENTS, by Roxana H. Vivian. The Graphic Press, Newton, Mass.

THE WOOD INDUSTRIES, by M. C. Mills. Regional plane of New York and its environs. Economic Series, Monograph No. 4.

THE GREATEST AND LEAST VARIATE UNDER GENERAL LAWS OF ERROR, by E. L. Dodd. Reprinted Translated from American Mathematical Society, Oct., 1923.

CONSUMPTION, MERCHANDISING, AND ADVERTISING OF FOODS, by A. E. Taylor. April, 1924. Reprinted from Harvard Business Review.

COÖPERATIVE PRODUCTION AND MARKETING, A. E. Taylor. Food Research Institute, Stanford University.

THE FOOD NEEDS OF THE NATION IN RELATION TO ECONOMIC RESOURCES, by A. E. Taylor. Address delivered before International Health Conference, San Francisco, July, 1923.

OTHER BOOKS

NATIONAL INDUSTRIAL CONFERENCE BOARD. *Legal restrictions on hours of work in the United States.* A reference manual. New York: 1924. v, 125 pp.

SELLS, ELIJAH W. *The natural business year and thirteen other themes.* Chicago, Shaw, 1924. xxviii, 276 pp.

CROWELL, SEYMOUR L. *New York Stock Exchange, Report of the President, May 1, 1923–May 1, 1924.* 31 pp.

HERMBERG, PAUL. *Der Kampf und der Weltmarket.* Jena, 1920. 135 pp.

BOLINE, ANTONIN. *La Capitale de Prague,* etude demographique. Prague, 1923. Knihovian Statistickeho Vestniku Svazeh 3. 99 pp.

NEW SERIES, NO. 148 (VOL. XIX) DECEMBER, 1924

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

CONTENTS

A METHOD OF HANDLING CURVILINEAR CORRELATION FOR ANY NUMBER OF VARIABLES. By Mordecai Ezekiel 431

A NEW ZEALAND STUDY IN NATALITY. By Edward P. Neale 454

PRINCIPLES UNDERLYING THE ISOLATION OF CYCLES AND TRENDS. By Willford I. King 468

GROWTH CURVES AND RAILWAY TRAFFIC. By Leroy E. Peabody 476

COST OF LIVING STATISTICS OF THE U. S. BUREAU OF LABOR STATISTICS AND THE NATIONAL INDUSTRIAL CONFERENCE BOARD. By Elma Carr 484

NOTES:

Three Years' Progress in Current Business Statistics. By Mortimer B. Lane 508

A Shortened Interpolation Formula for Certain Types of Data. By E. P. Neale and D. M. Y. Somerville 515

Wage Indexes and Price Indexes. By George Soule 517

Sex Cleavage in the Vote of 1920 (519); Quetelet's Anniversary (520); Welfare Activities of the U. S. Steel Corporation (522); Progress of Work in the Census Bureau (524); Miscellaneous Notes (525).

REVIEWS: Viner: *Canada's Balance of International Indebtedness*, 1900–1913, by R. H. Coats (534); Cohen: *Insurance by Industry*, by B. M. Stewart (540); Wright: *Sugar in Relation to the Tariff*, by J. Viner (544); Willis and Byers: *Portland Cement Prices*, by I. Lubin (547); Downey: *Workmen's Compensation*, by M. M. Dawson (550); Stella: *Some Aspects of Italian Immigration to the United States*, by R. F. Foerster (553); Monroe: *Life of Francis A. Walker*, by E. W. Kopf (554); Powell: *Payment by Results*, by H. W. Rauber (555)

RECENT LITERATURE 557

Published quarterly by the American Statistical Association

Rumford Press Building, Concord, N. H.

Editorial Office, New York City

Entered at the post-office, Concord, N. H., as second-class mail matter

Acceptance for mailing at the special rate of postage provided for in Section No. 1103, Act of October 3, 1917, authorized April 29, 1922

THE RUMFORD PRESS
CONCORD N. H.